INTRODUCTION TO FREE-ENERGY CALCULATIONS

Chris Chipot

Laboratoire International Associé CNRS-UIUC,
Unité Mixte de Recherche n° 7565, Université de Lorraine

Beckman Institute for Advanced Science and Technology,
Department of Physics
University of Illinois at Urbana-Champaign

INTRODUCTION

The race for longer and larger simulations
What is the best method for a given problem?

ALCHEMICAL FREE-ENERGY CALCULATIONS

- A tool to address host-guest chemistry questions
- Good practices, guidelines and recommendations
- The long-standing protein-ligand problem

GEOMETRICAL FREE-ENERGY CALCULATIONS

- What is a good reaction-coordinate model?
- A host of methods to measure free-energy changes
- Potentials of mean force and transport phenomena
- Potentials of mean force and recognition and association phenomena
- What about non-equilibrium work computer experiments?

ONGOING CHALLENGES AT THE FRONTIERS OF FREE-ENERGY CALCULATIONS

CONCLUDING REMARKS AND QUESTIONS

INTRODUCTION

The race for longer and larger simulations
What is the best method for a given problem?

ALCHEMICAL FREE-ENERGY CALCULATIONS

- A tool to address host-guest chemistry questions
- Good practices, guidelines and recommendations
- The long-standing protein-ligand problem

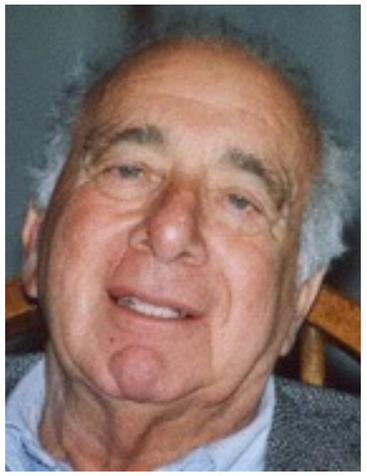
GEOMETRICAL FREE-ENERGY CALCULATIONS

- What is a good reaction-coordinate model?
- A host of methods to measure free-energy changes
- Potentials of mean force and transport phenomena
- Potentials of mean force and recognition and association phenomena
- What about non-equilibrium work computer experiments?

ONGOING CHALLENGES AT THE FRONTIERS OF FREE-ENERGY CALCULATIONS

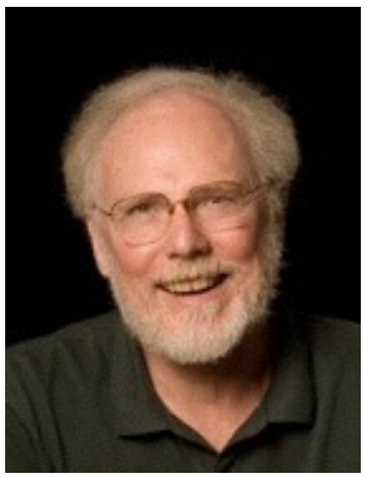
CONCLUDING REMARKS AND QUESTIONS

THE RACE FOR LONGER AND LARGER SIMULATIONS



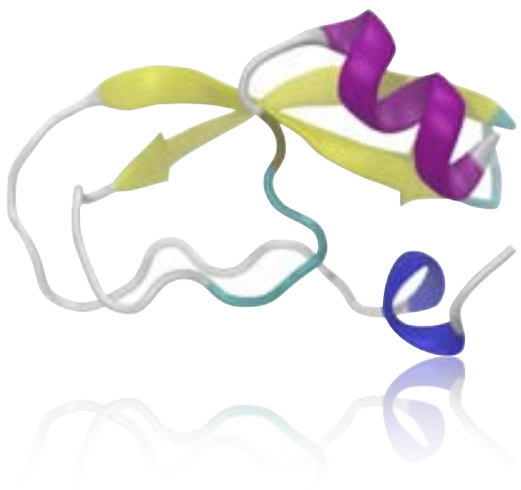
First molecular dynamics simulation. Phase transition in model liquids.

$$\begin{cases} m_i \dfrac{\mathrm{d}^2 x_i}{\mathrm{d}t^2} &= F_i \\ F_i &= -\dfrac{\partial U(\mathbf{x})}{\partial x_i} \end{cases}$$



A TURNING POINT IN COMPUTATIONAL STRUCTURAL BIOLOGY

First molecular dynamics simulation applied to a small protein, BPTI, over 8 ps.



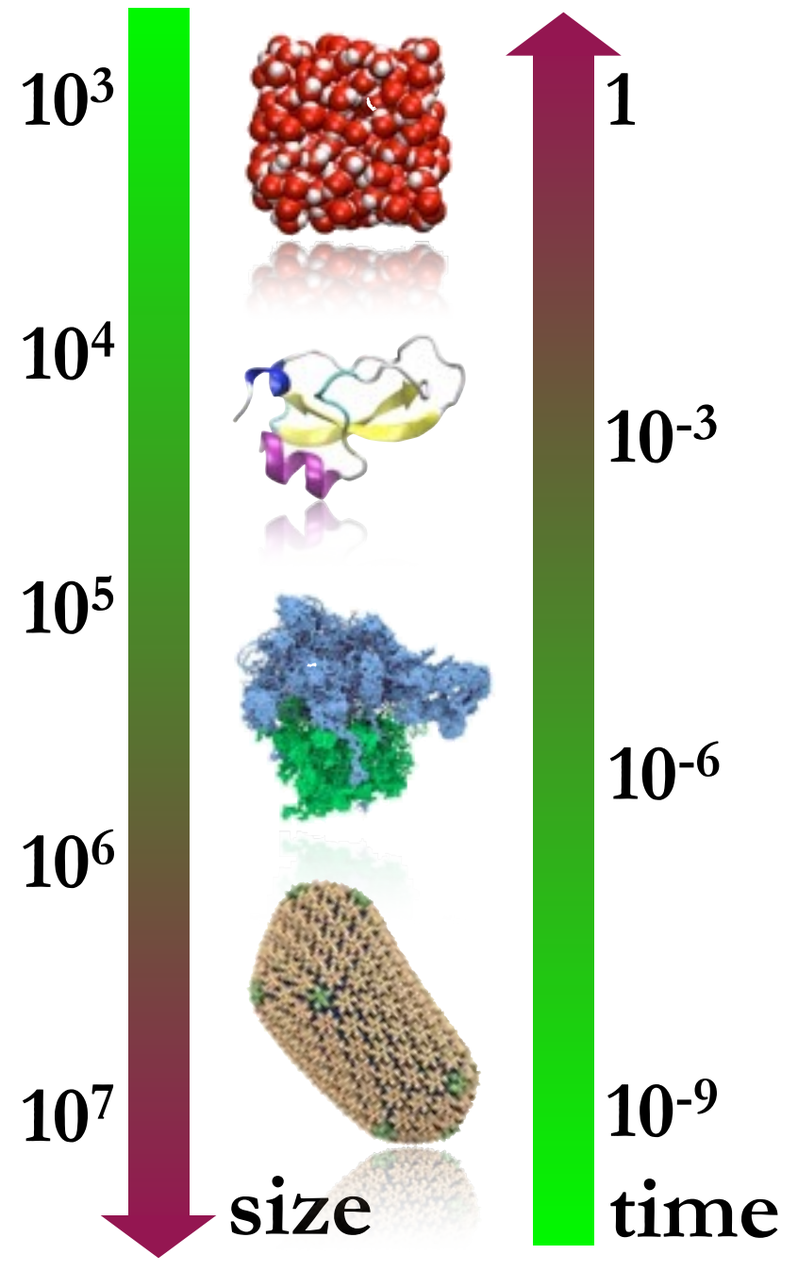
Alder, B. J.; Wainwright, T. E. *J. Chem. Phys.*, **1957**, *27*, 1208-1209

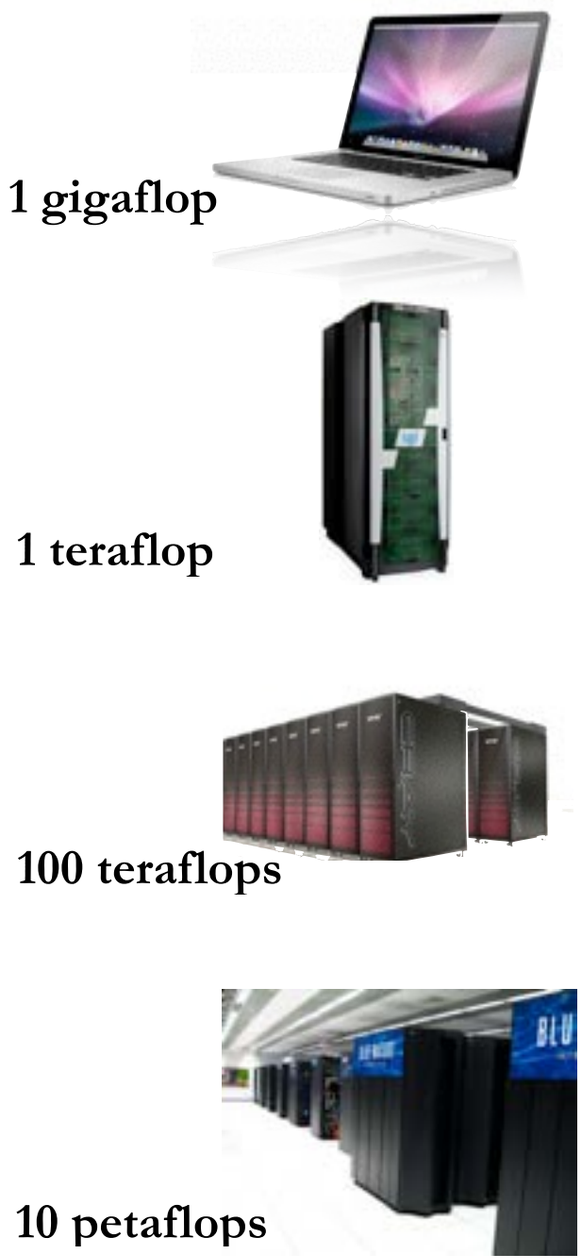
McCammon, J. A.; Gelin, B. R.; Karplus, M. *Nature,* **1977**, *267*, 585-590

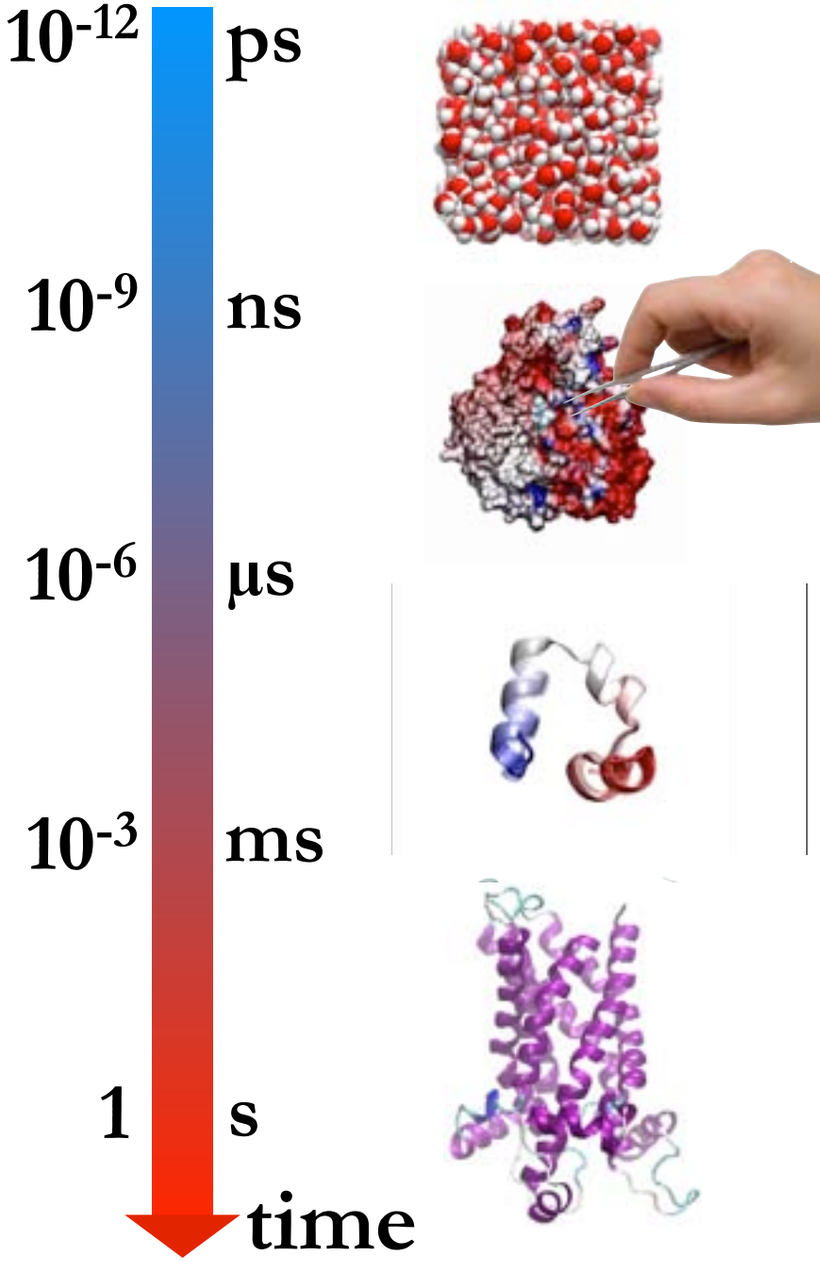
THE RACE FOR LONGER AND LARGER SIMULATIONS

Brute-force simulations struggling to bridge the gap between time and size scales

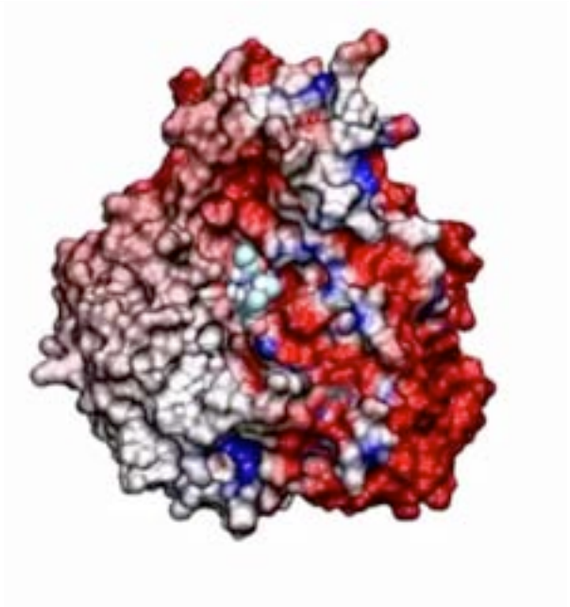
Free-energy calculations are computational tweezers to anatomize and explore relevant degrees of freedom



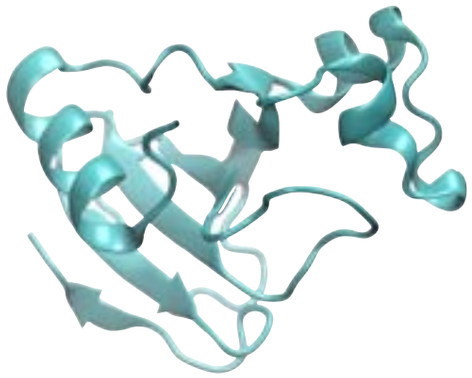


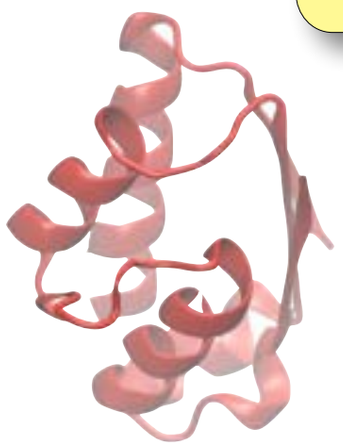


WHAT ARE FREE-ENERGY CALCULATIONS COMMONLY USED FOR ?

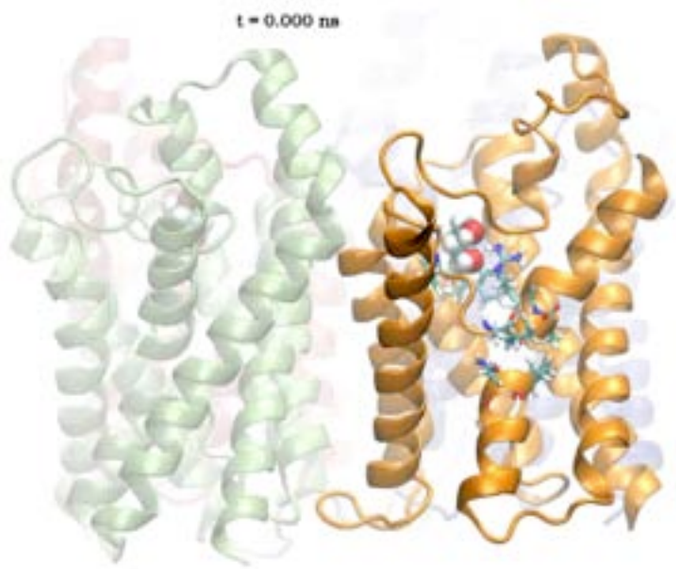


Recognition and association phenomena

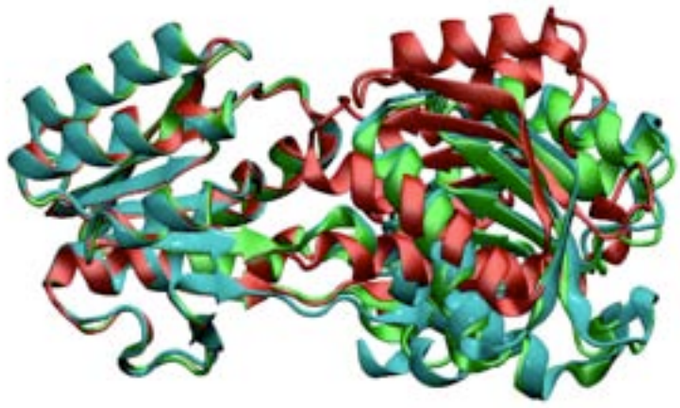




- protein-ligand binding
- site-directed mutagenesis
- protein-protein binding
- partition coefficients
- permeabilities
- activation barriers
-structural modifications



Transport phenomena

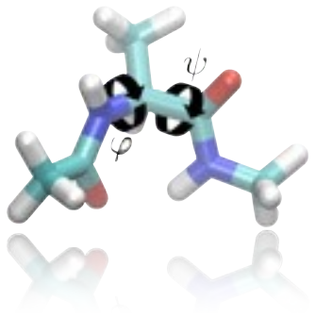


Conformational transitions

WHAT IS THE BEST METHOD FOR A GIVEN PROBLEM ?

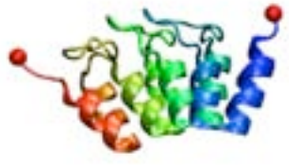
Free-energy differences can be estimated computationally following four possible routes

(1) Methods based on histograms



$$\Delta A(\xi) = -\frac{1}{\beta} \ln P(\xi) + \Delta A_0$$

(2) Non-equilibrium work simulations

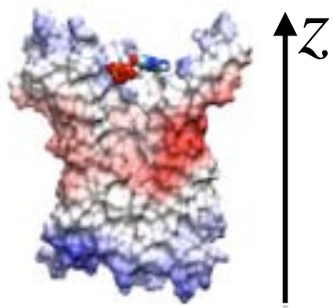


$$\exp(-\beta \Delta A) = \langle \exp(-\beta w) \rangle$$

(3) Perturbation theory

$$\exp(-\beta \Delta A) = \langle \exp(-\beta \Delta U) \rangle_0$$

(4) Measuring the derivative and integrating it



$$\frac{\mathrm{d}A(\xi)}{\mathrm{d}\xi} = \left\langle \frac{\partial U}{\partial \xi} - \frac{1}{\beta} \frac{\partial \ln |J|}{\partial \xi} \right\rangle_\xi$$

Torrie, G. M.; Valleau, J. P. *Chem. Phys. Lett.* **1974**, *28*, 578-581

Widom, B. *J. Chem. Phys.* **1963**, *39*, 2808-2812

Israelewitz, B.; Gao, M.; Schulten, K. *Curr. Opin. Struct. Biol.* **2001**, *11*, 224-230

Jarzynski, C. *Phys. Rev. Lett.* **1997**, *78*, 2690-2693

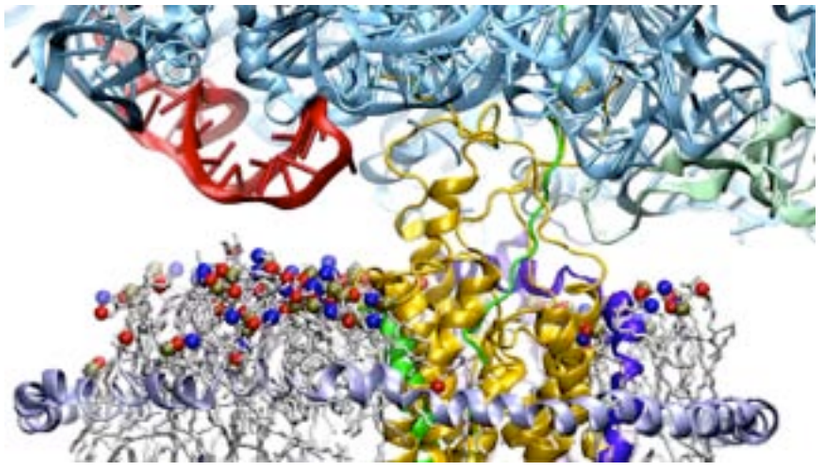
Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420-1426

Pohorille, A.; Jarzynski, C.; Chipot, C. *J. Phys. Chem. B* **2010**, *114*, 10235-10253

Kirkwood, J. G. *J. Chem. Phys.* **1935**, *3*, 300-313

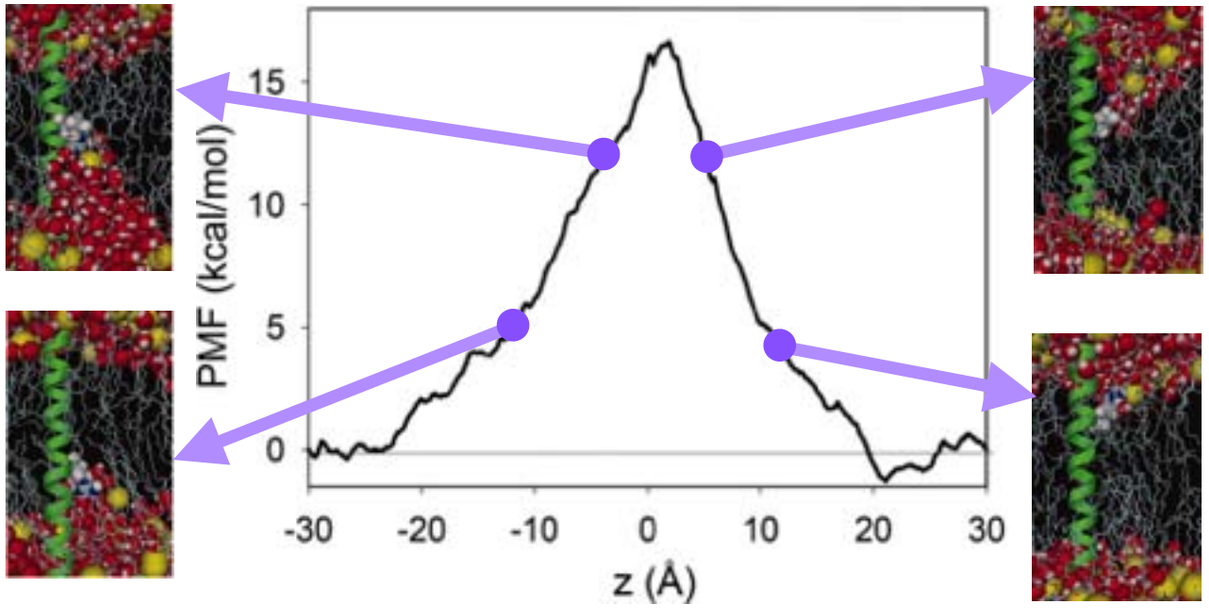
Carter, E. et al. *Chem. Phys. Lett.* **1989**, *156*, 472-477

WHAT IS THE BEST METHOD FOR A GIVEN PROBLEM ?

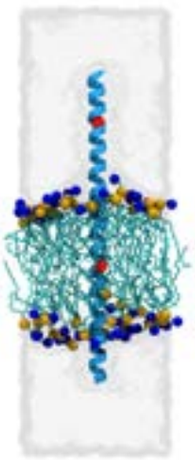


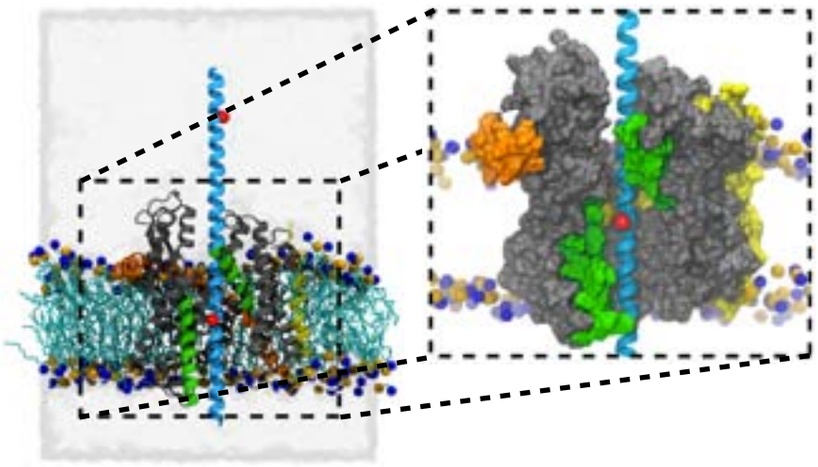
Nascent membrane proteins typically insert into the membrane via the Sec-translocon.

Membrane insertion of arginine, for instance, requires 14–17 kcal/mol according to molecular dynamics simulations, but only 2–3 kcal/mol according to experiment.



How does the translocon reduce the energetic cost and gain that accompanies insertion?





Becker, T. et al. *Science* **2009**, *326*, 1369-1373

Dorairaj, S.; Allen, T. W. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 4943-4948

Hessa, T. et al. *Nature* **2007**, *450*, 1026-1030

Gumbart, J. C.; Chipot, C.; Schulten, K. *Proc. Natl Acad. Sci. USA* **2011**, *108*, 3596-3601

INTRODUCTION

The race for longer and larger simulations
What is the best method for a given problem?

ALCHEMICAL FREE-ENERGY CALCULATIONS

- A tool to address host-guest chemistry questions
- Good practices, guidelines and recommendations
- The long-standing protein-ligand problem

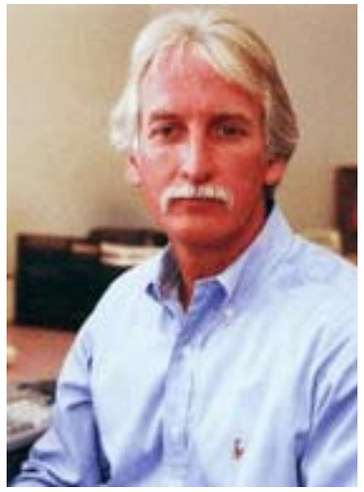
GEOMETRICAL FREE-ENERGY CALCULATIONS

- What is a good reaction-coordinate model?
- A host of methods to measure free-energy changes
- Potentials of mean force and transport phenomena
- Potentials of mean force and recognition and association phenomena
- What about non-equilibrium work computer experiments?

ONGOING CHALLENGES AT THE FRONTIERS OF FREE-ENERGY CALCULATIONS

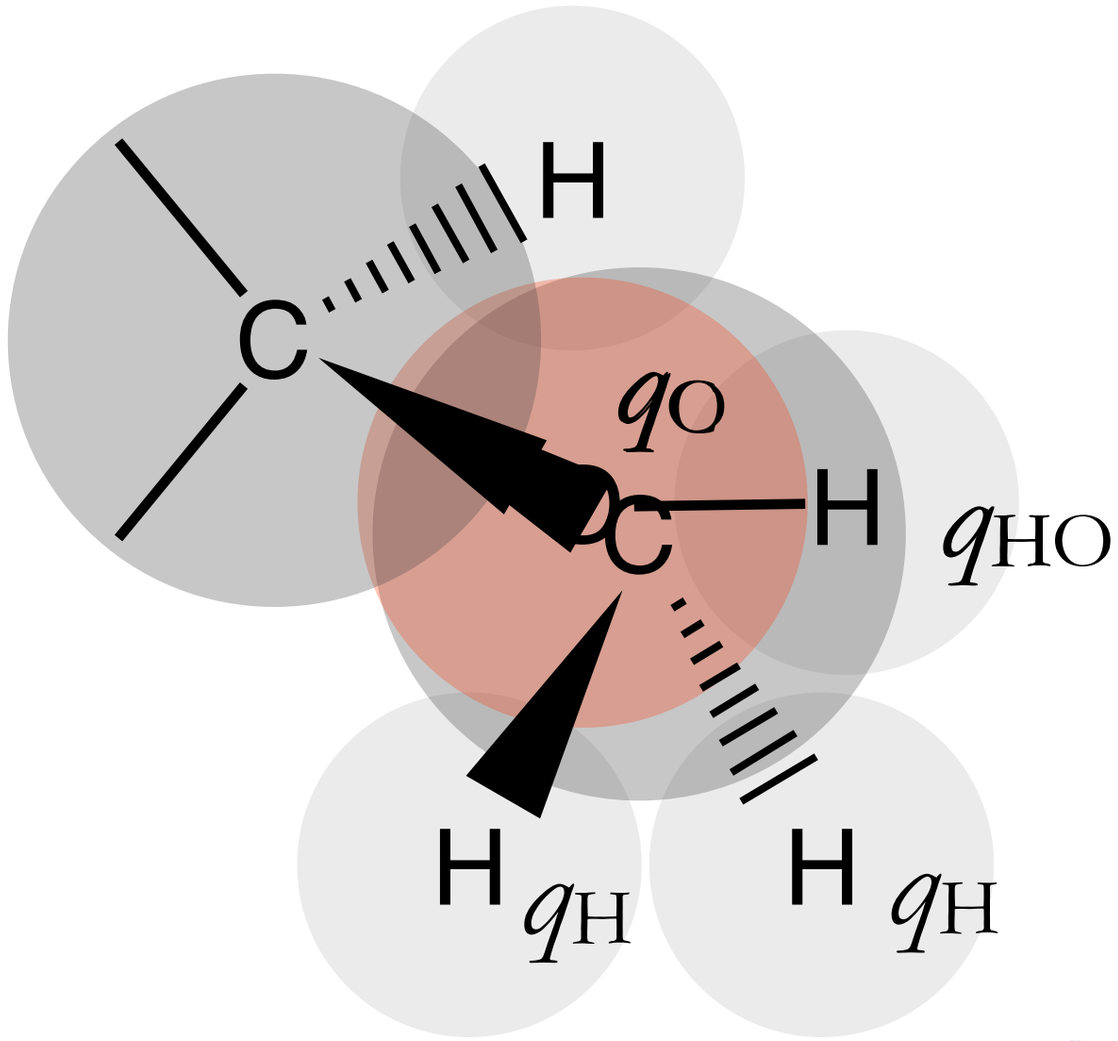
CONCLUDING REMARKS AND QUESTIONS

A TOOL TO ADDRESS HOST-GUEST CHEMISTRY PROBLEMS



Transforming between chemical species, exploiting the malleability of the potential energy function.

First alchemical transformation: Methanol to ethane (6.7 vs. 6.9 kcal/mol in experiment).



Single-topology paradigm:

- A common-denominator topology for the end states of the transformation.

- The non-bonded terms are a function of a general-extent parameter, λ.

- Necessity to correct for the change of bond length.

- Requires electrostatic decoupling.

Jorgensen, W. L.; Ravimohan, C. *J. Chem. Phys.* **1985**, *83*, 3050-3054

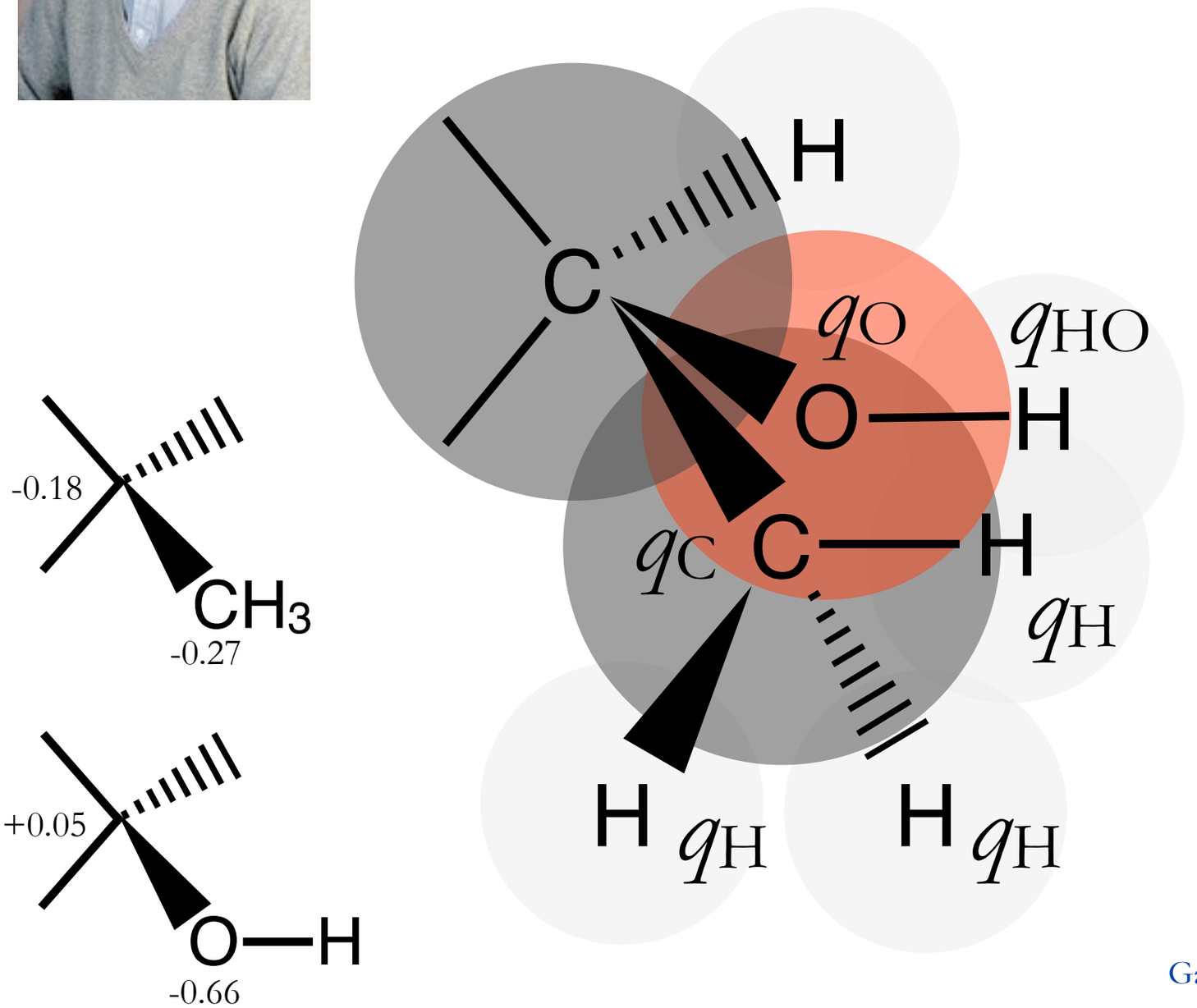
Bash, P. A. et al. *Science* **1987**, *236*, 564-568

Bash, P. A. et al. *Science* **1987**, *235*, 574-576

A TOOL TO ADDRESS HOST-GUEST CHEMISTRY PROBLEMS

Free energy is a state function.

The free-energy difference between the end states is independent from the path followed to calculate it.

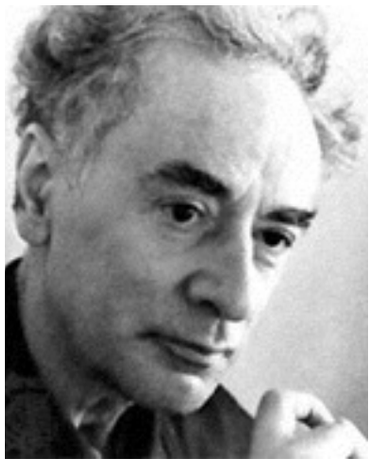


Dual-topology paradigm:

- The end states are defined explicitly.
- An exclusion list prevents them from seeing each other.
- The interaction of the perturbed state with its environment is a function of a general-extent parameter, λ.
- Prone to end-point catastrophes.
- Branching requires particular care.

Gao, J.; Kuczera, K.; Tidor, B.; Karplus, M. *Science* **1989**, *244*, 1069-1072

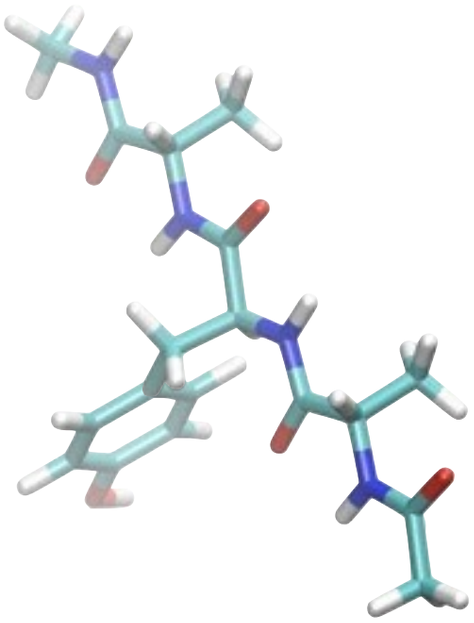
A TOOL TO ADDRESS HOST-GUEST CHEMISTRY PROBLEMS



FREE-ENERGY PERTURBATION

$$\exp(-\beta\Delta A) = \langle \exp(-\beta\Delta U) \rangle_0$$

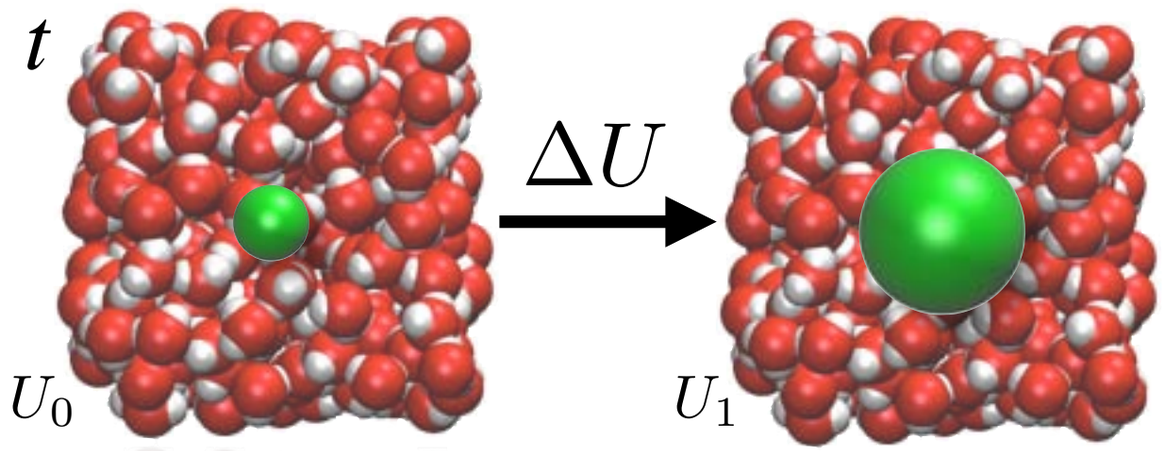
- Requires the sole knowledge of the reference state.
- Formally exact for any perturbation.
- Importance-sampling method.



$$\Delta U(\lambda) = \lambda U_1 + (1-\lambda)U_0$$

THERMODYNAMIC INTEGRATION

$$\Delta A = \int \mathrm{d}\lambda \left\langle \frac{\partial U}{\partial \lambda} \right\rangle_\lambda$$



Landau, L. D. Statistical physics, **1938**

Zwanzig, R. W. J. *Chem. Phys.* **1954**, *22*, 1420-1426

Kirkwood, J. G. *J. Chem. Phys.* **1935**, *3*, 300-313

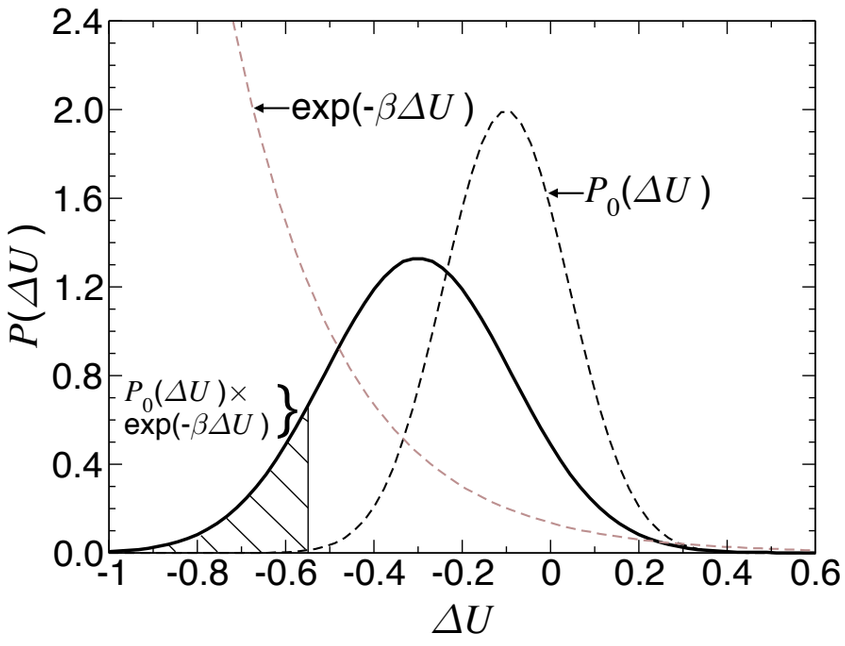
GOOD PRACTICES, GUIDELINES AND RECOMMENDATIONS

How to deal with large perturbations ?

$$\Delta A = -\frac{1}{\beta} \ln \int \mathrm{d}\Delta U \; P_0(\Delta U) \; \exp(-\beta \Delta U)$$

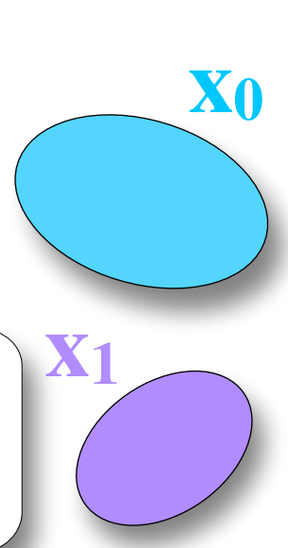
$$P_0(\Delta U) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \; \exp\left[-\frac{(\Delta U - \langle \Delta U \rangle_0)^2}{2\sigma_0^2}\right]$$

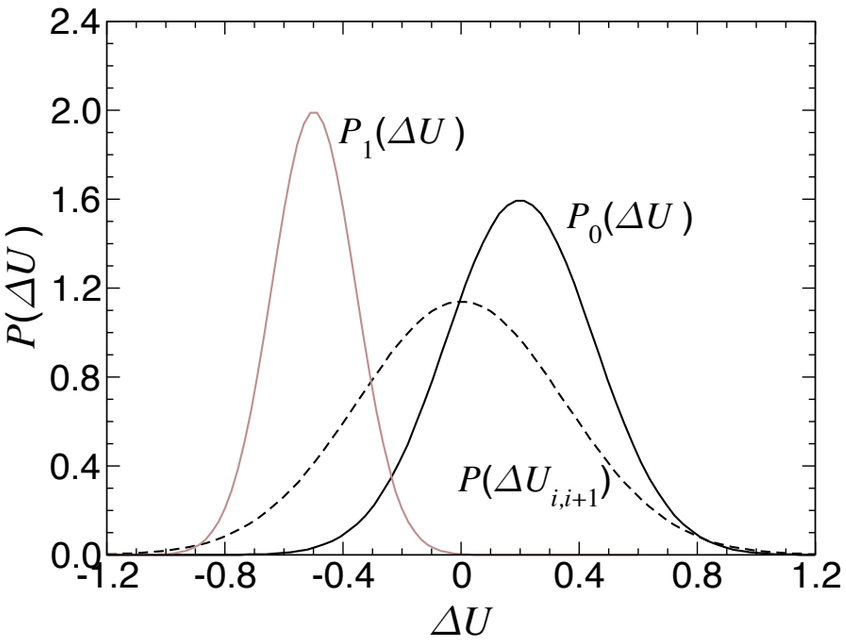
$$\Delta A = \langle \Delta U \rangle_0 - \frac{1}{2}\beta\sigma_0^2$$



Stratification strategies

$$\Delta A = -\frac{1}{\beta} \sum_i \ln \langle \exp(-\beta \Delta U_{i,i+1}) \rangle_i$$





Valleau, J. P.; Card, D. N. *J. Chem. Phys.* **1972**, *57*, 5457-5462

Chipot, C.; Pohorille, A. Free energy calculations. Theory and applications in chemistry and biology, **2007**

Lelièvre, T.; Stoltz, G.; Rousset, M. Free energy computations: A mathematical perspective, **2010**

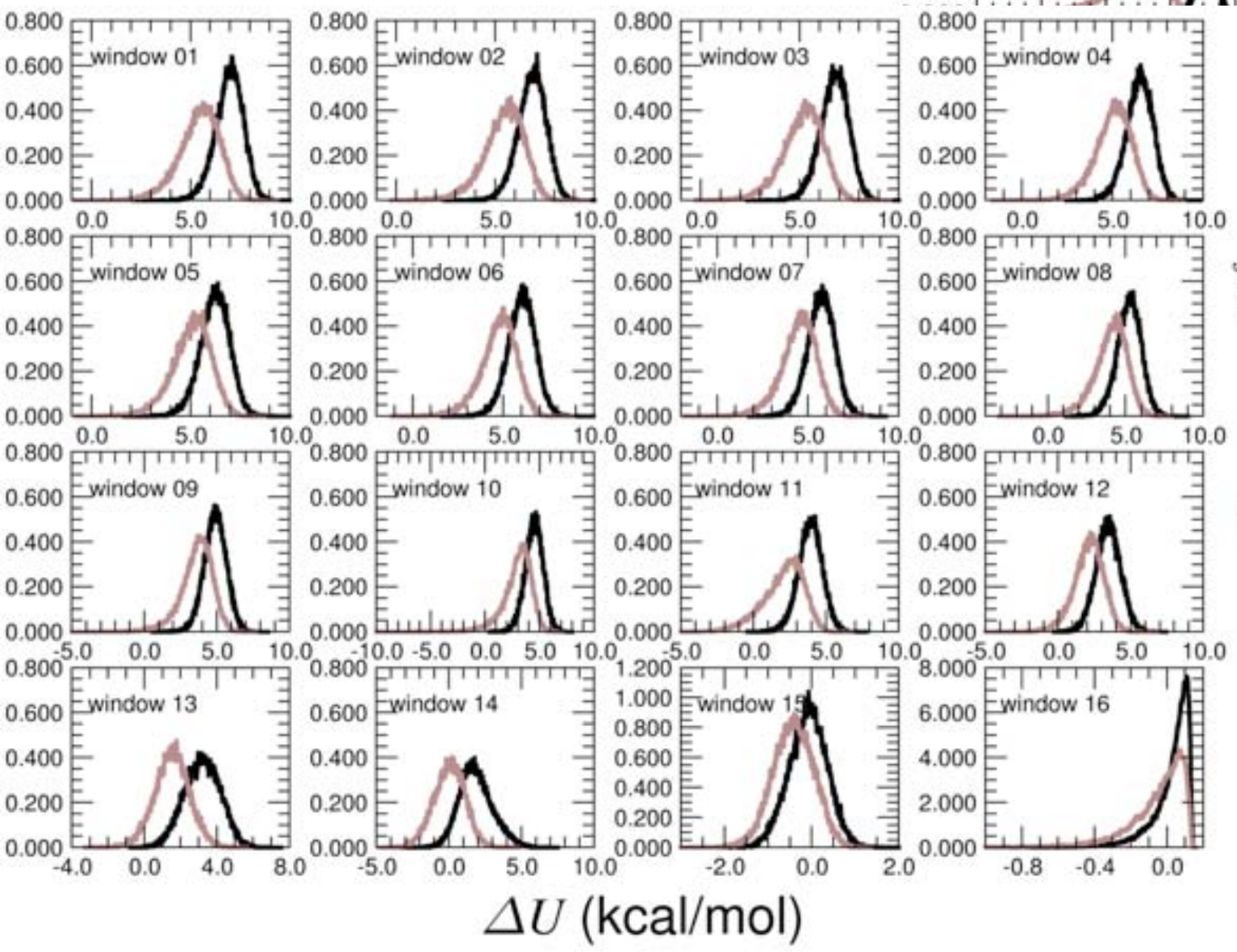
GOOD PRACTICES, GUIDELINES AND RECOMMENDATIONS

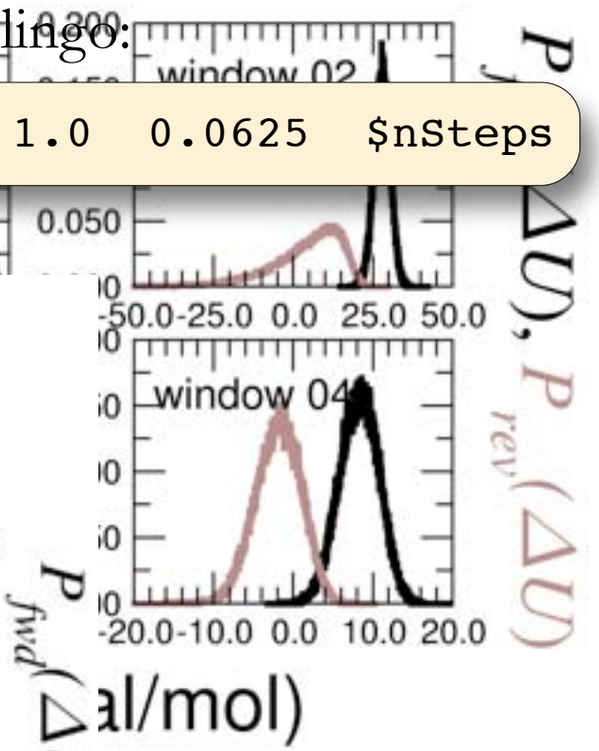
How many strata should I choose ?

Stratification will impact the accuracy of the free-energy calculation.

In the NAMD lingo:

```
runFEP   0.0   1.0   0.0625   $nSteps
```

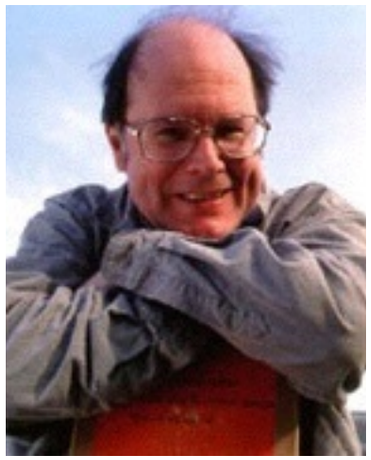




Valleau, J. P.; Card, D. N. *J. Chem. Phys.* **1972**, *57*, 5457-5462

Pohorille, A.; Jarzynski, C.; Chipot, C. *J. Phys. Chem. B* **2010**, *114*, 10235-10253

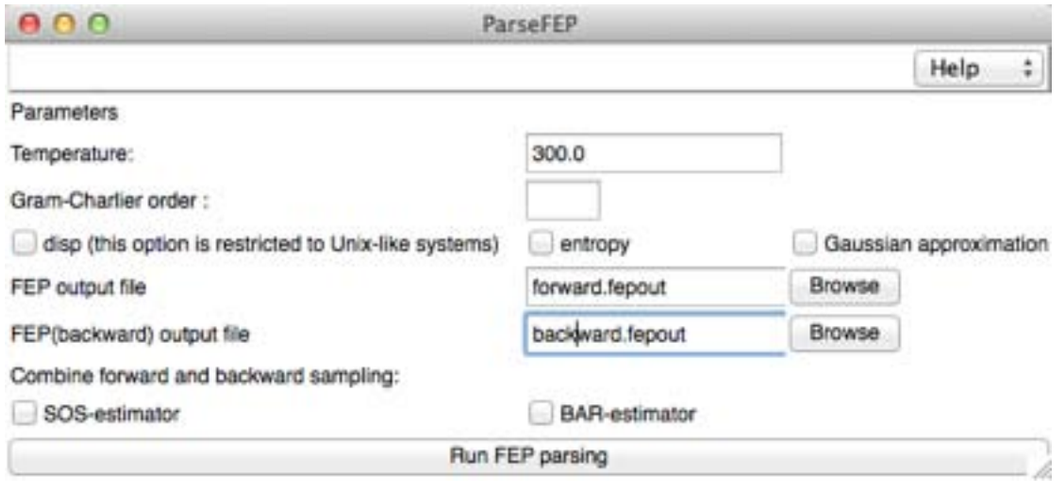
GOOD PRACTICES, GUIDELINES AND RECOMMENDATIONS



Combining forward and backward tranformations

Maximum-likelihood estimator of the free-energy change.

Guarantees the minimum variance.



$$\begin{cases} \exp\left(\beta\Delta\hat{A}^{\mathrm{BAR}}\right) = \dfrac{\langle f\left[-\beta\left(\Delta U - c\right)\right]\rangle_1}{\langle f\left[+\beta\left(\Delta U - c\right)\right]\rangle_0}\ \exp\left(+\beta c\right) \\ c = \Delta\hat{A}^{\mathrm{BAR}} + \dfrac{1}{\beta}\ln\dfrac{N_1}{N_0} \qquad f(x) = 1/[1+\exp(x)] \end{cases}$$

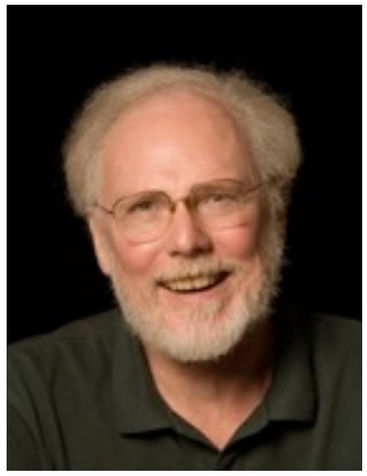
$$\sigma^2_{\Delta A}{}^{\mathrm{BAR}} = \frac{1}{N_0\beta^2}\left[\frac{\langle f^2(x)\rangle_0}{\langle f(x)\rangle_0^2} - 1\right] + \frac{1}{N_1\beta^2}\left[\frac{\langle f^2(-x)\rangle_1}{\langle f(-x)\rangle_1^2} - 1\right]$$

Bennett, C. H. *J. Comp. Phys.* **1976**, *22*, 245–268.

Pohorille, A.; Jarzynski, C.; Chipot, C. *J. Phys. Chem. B* **2010**, *114*, 10235-10253

Hahn, A. M.; Then, H. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **2009**, *80*, 031111

GOOD PRACTICES, GUIDELINES AND RECOMMENDATIONS

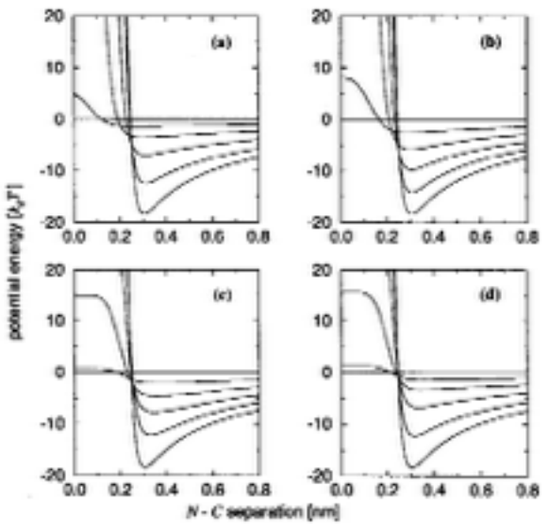


What about end-point catastrophes ?

Avoid singularities in the van der Waals potential as particles appear.

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

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



In the NAMD lingo:

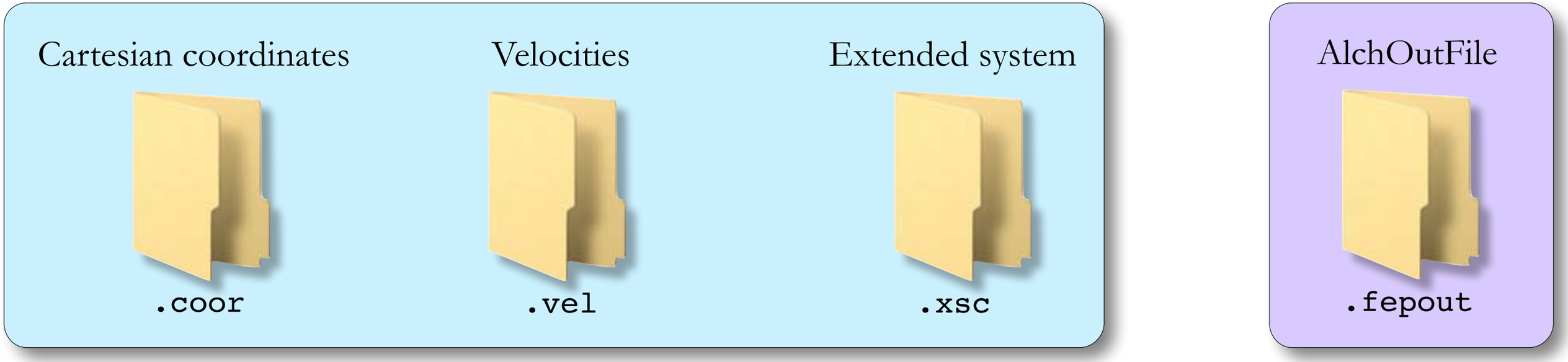
```
alchVdWShiftCoeff     4.0
```

Zacharias, M.; Straatsma, T. P.; McCammon, J. A. *J. Chem. Phys.* **1994**, *100*, 9025-9031

Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. *Chem. Phys. Lett.* **1994**, *222*, 529-539

GOOD PRACTICES, GUIDELINES AND RECOMMENDATIONS

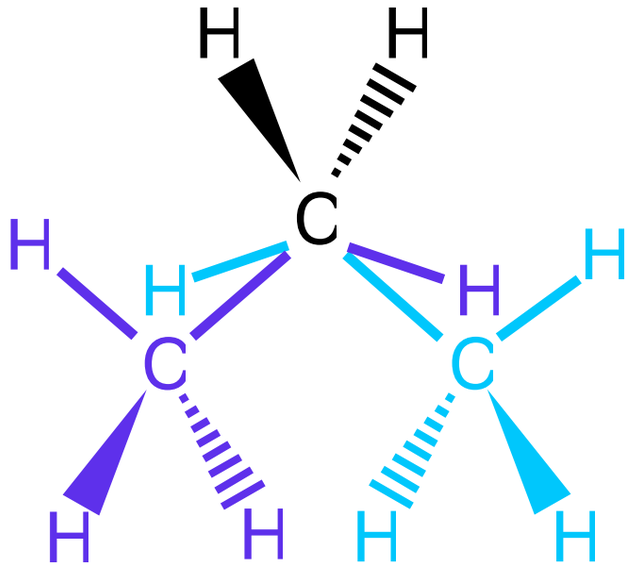
Equilibration simulation



Phillips, J. C. et al. *J. Comput. Chem.* **2005**, *26*, 1781-1802

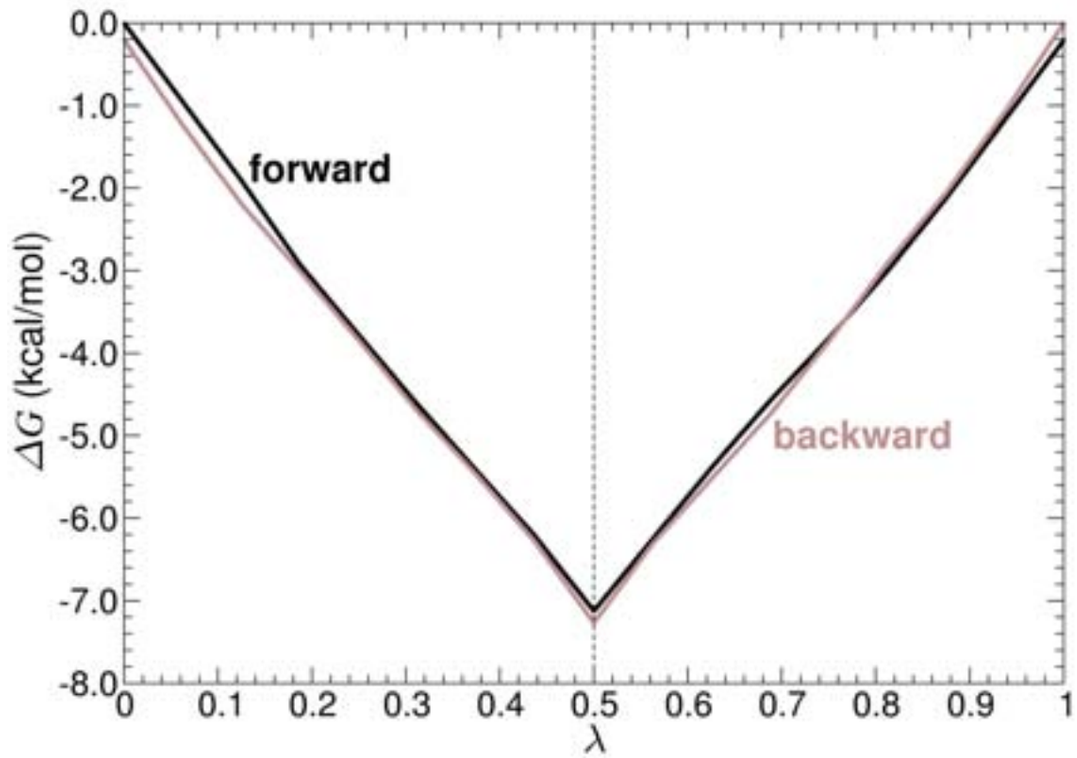
Bhandarkar, M. et al. NAMD user's guide, version 2.9, **2012**

GOOD PRACTICES, GUIDELINES AND RECOMMENDATIONS




```
RESI ZERO          0.00       !  ethane -> ethane
GROUP                         !
ATOM CI    CT3    -0.27       !
ATOM HI1   HA      0.09       !
ATOM HI2   HA      0.09       !
ATOM HI3   HA      0.09       !
GROUP                         !   HI1        HM1    HF2  HF3
ATOM CM    CT3    -0.27       !      \         |      |  /
ATOM HM1   HA      0.09       !       \HF      |      | /
ATOM HM2   HA      0.09       !        CI-----CM-----CF
ATOM HI    HA      0.09       !       / |      |      HI\
ATOM HF    HA      0.09       !      /  |      |         \
GROUP                         !   HI2  HI3   HM2         HF1
ATOM CF    CT3    -0.27       !
ATOM HF1   HA      0.09       !
ATOM HF2   HA      0.09       !
ATOM HF3   HA      0.09       !
BOND      CI   HI1        CI   HI2        CI    HI3        ! ethane 1
BOND      CF   HF1        CF   HF2        CF    HF3        ! ethane 2
BOND      CI   CM         CF   CM                           ! common
BOND      CM   HM1        CM   HM2                          ! common
BOND      CM   HI                                           ! ethane 1
BOND      CM   HF                                           ! ethane 2
```

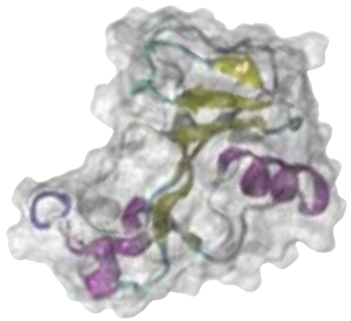
Zero free-energy change transformation



Decoupling in the NAMD lingo:

```
alchVdwLambdaEnd          1.0
alchElecLambdaStart       0.5
```

Pearlman, D. A.; Kollman, P. A. *J. Chem. Phys.* **1991**, *94*, 4532-4545



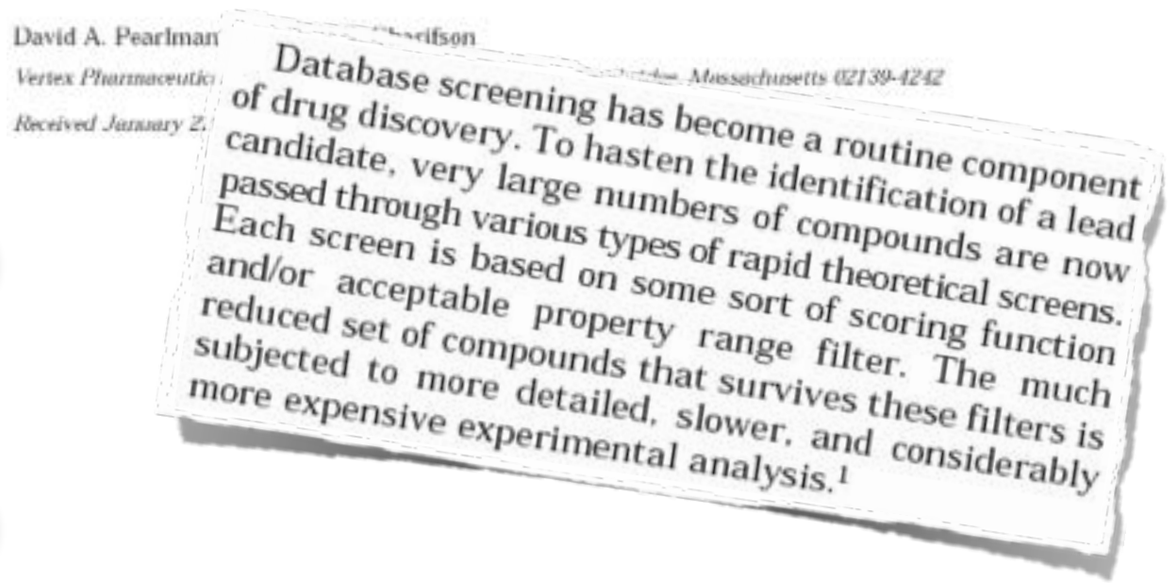
constant computational investment

shape recognition

rigid docking

flexible docking

free energy



J. Med. Chem. 2001, 44, 3417–3423 3417

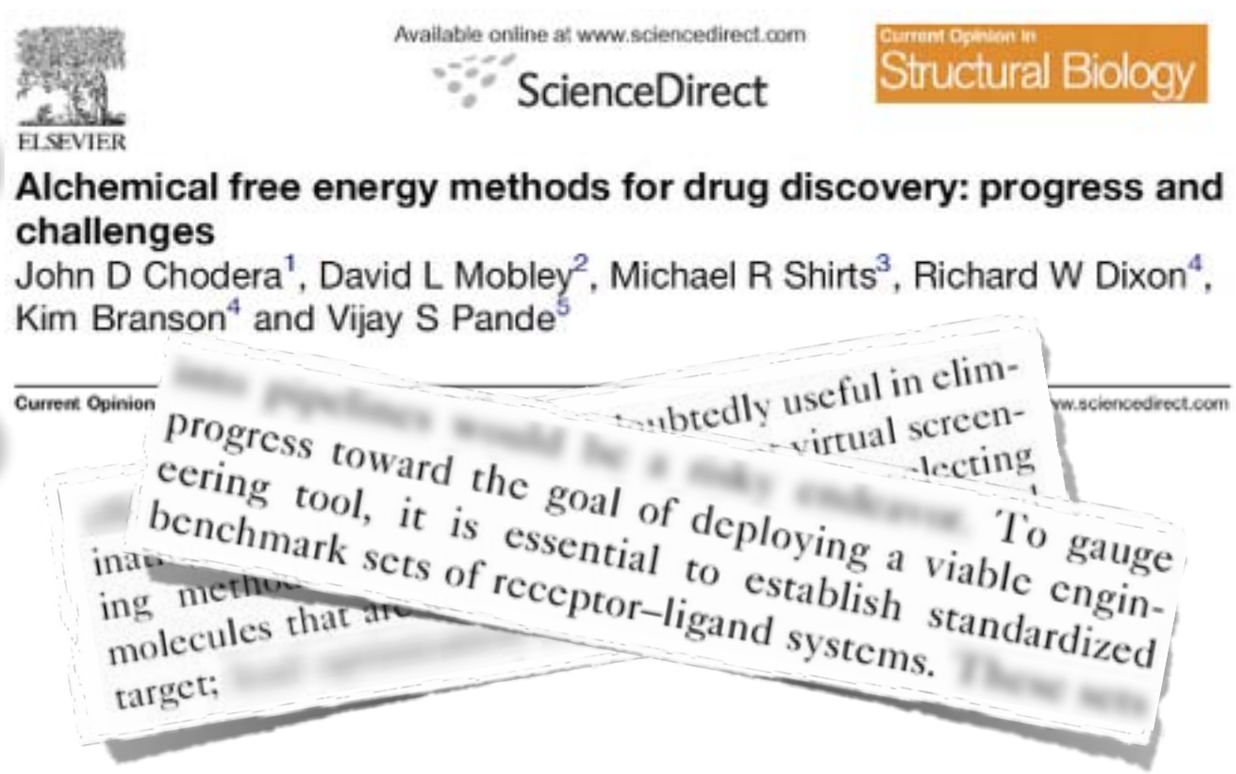
Are Free Energy Calculations Useful in Practice? A Comparison with Rapid Scoring Functions for the p38 MAP Kinase Protein System†

David A. Pearlman ... Sherifson

Vertex Pharmaceutic... Massachusetts 02139-4242

Received January 2...

Database screening has become a routine component of drug discovery. To hasten the identification of a lead candidate, very large numbers of compounds are now passed through various types of rapid theoretical screens. Each screen is based on some sort of scoring function and/or acceptable property range filter. The much reduced set of compounds that survives these filters is subjected to more detailed, slower, and considerably more expensive experimental analysis.[1]



Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

ELSEVIER

Alchemical free energy methods for drug discovery: progress and challenges

John D Chodera[1], David L Mobley[2], Michael R Shirts[3], Richard W Dixon[4], Kim Branson[4] and Vijay S Pande[5]

Current Opinion ... www.sciencedirect.com

...btedly useful in elim- ...virtual screen- ...lecting

progress toward the goal of deploying a viable engineering tool, it is essential to establish standardized benchmark sets of receptor–ligand systems.

To gauge

ing meth...

molecules that ar...

target;

Chipot, C.; Rozanska, X.; Dixit, S. B. *J. Comput. Aided Mol. Des.* **2005**, *19*, 765-770.

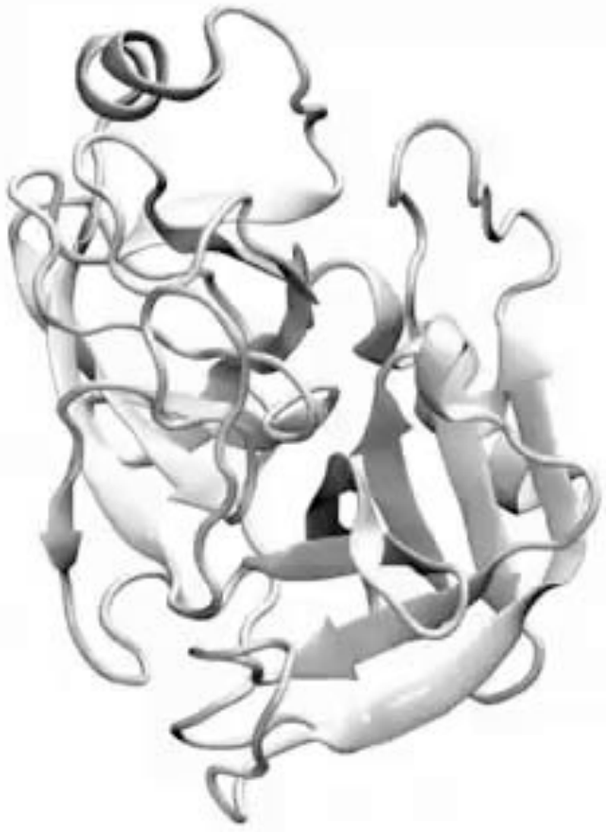
Shirts, M. R.; Mobley, D. L.; Chodera, J. D. *Annual Reports Comput. Chem.* **2007**, 3, 41-59.

Chipot, C. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4*, 71-89.

THE LONG-STANDING PROTEIN-LIGAND PROBLEM

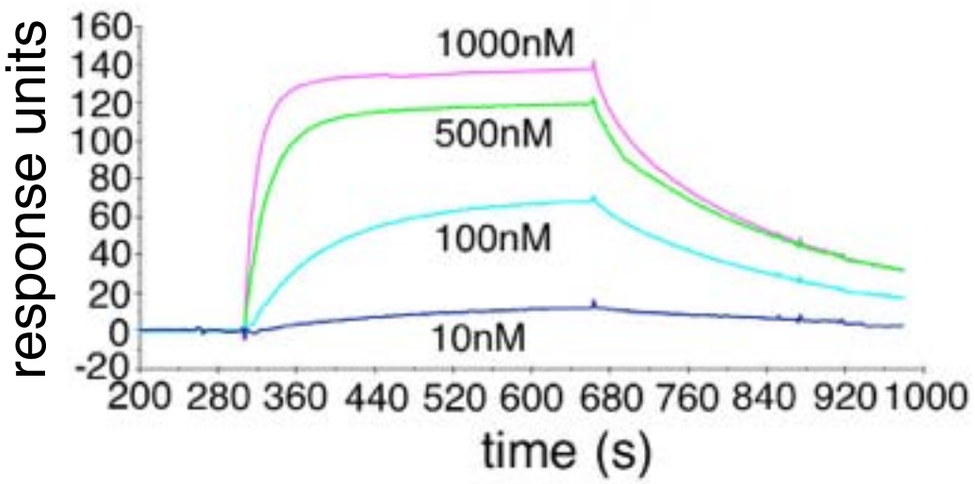
$$K_{\text{eq}} = \frac{[\text{protein : ligand}]}{[\text{protein}][\text{ligand}]}$$

$$\text{protein} + \text{ligand} \overset{K_{\text{eq}}}{\rightleftharpoons} \text{protein : ligand}$$



which can readily be determined by experiment:

$$K_d = \frac{k_{\text{off}}}{k_{\text{on}}}$$

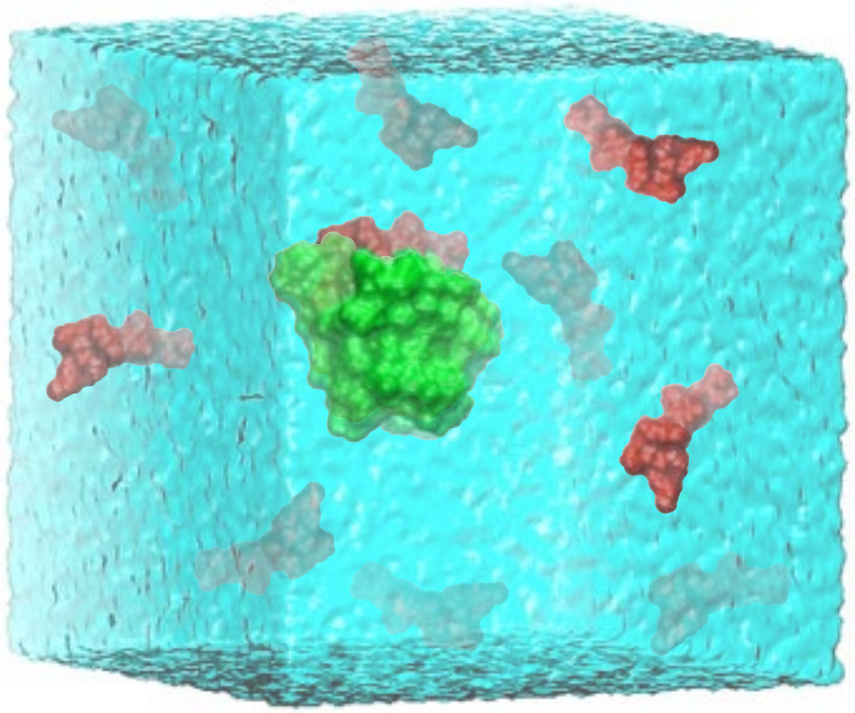


— A single event is evidently not enough.

— Brute-force simulations are limited by k_{on} and k_{off}.

Kollman, P.A. *Chem. Rev.* **1993**, *93*, 2395-2417
Gilson, M. K. et al. *Biophys. J.* **1997**, *72*, 1047-1069
Chipot, C.; Pohorille, A. Free-energy calculations. Springer **2007**.
Karlsson, R.; Larsson, A. *Methods Mol. Biol.* **2004**, *248*, 389-415
Buch, I.; Giorgino, T.; Fabritiis, G. D. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 10184-10189

The long-standing protein-ligand problem



$$K_{\text{eq}} = \frac{[\text{protein} : \text{ligand}]}{[\text{protein}][\text{ligand}]}$$

$$\text{protein} + \text{ligand} \overset{K_{\text{eq}}}{\rightleftharpoons} \text{protein} : \text{ligand}$$

$$K_{\text{eq}} = \frac{p_1 [\text{protein}]_{\text{tot}}}{[\text{ligand}]\ p_0\ [\text{protein}]_{\text{tot}}} = \frac{1}{[\text{ligand}]}\ \frac{p_1}{p_0}$$

$$K_{\text{eq}} = \frac{1}{[\text{ligand}]} \left\{ \frac{\int_{\text{site}} \mathrm{d}\mathbf{1} \int_{\text{bulk}} \mathrm{d}\mathbf{2} \ldots \int_{\text{bulk}} \mathrm{d}\mathbf{N} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta U}}{\int_{\text{bulk}} \mathrm{d}\mathbf{1} \int_{\text{bulk}} \mathrm{d}\mathbf{2} \ldots \int_{\text{bulk}} \mathrm{d}\mathbf{N} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta U}} \right.$$

$$+ \frac{\int_{\text{bulk}} \mathrm{d}\mathbf{1} \int_{\text{site}} \mathrm{d}\mathbf{2} \ldots \int_{\text{bulk}} \mathrm{d}\mathbf{N} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta U}}{\int_{\text{bulk}} \mathrm{d}\mathbf{1} \int_{\text{bulk}} \mathrm{d}\mathbf{2} \ldots \int_{\text{bulk}} \mathrm{d}\mathbf{N} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta U}} + \ldots$$

$$\left. + \frac{\int_{\text{bulk}} \mathrm{d}\mathbf{1} \int_{\text{bulk}} \mathrm{d}\mathbf{2} \ldots \int_{\text{site}} \mathrm{d}\mathbf{N} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta U}}{\int_{\text{bulk}} \mathrm{d}\mathbf{1} \int_{\text{bulk}} \mathrm{d}\mathbf{2} \ldots \int_{\text{bulk}} \mathrm{d}\mathbf{N} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta U}} \right\}$$

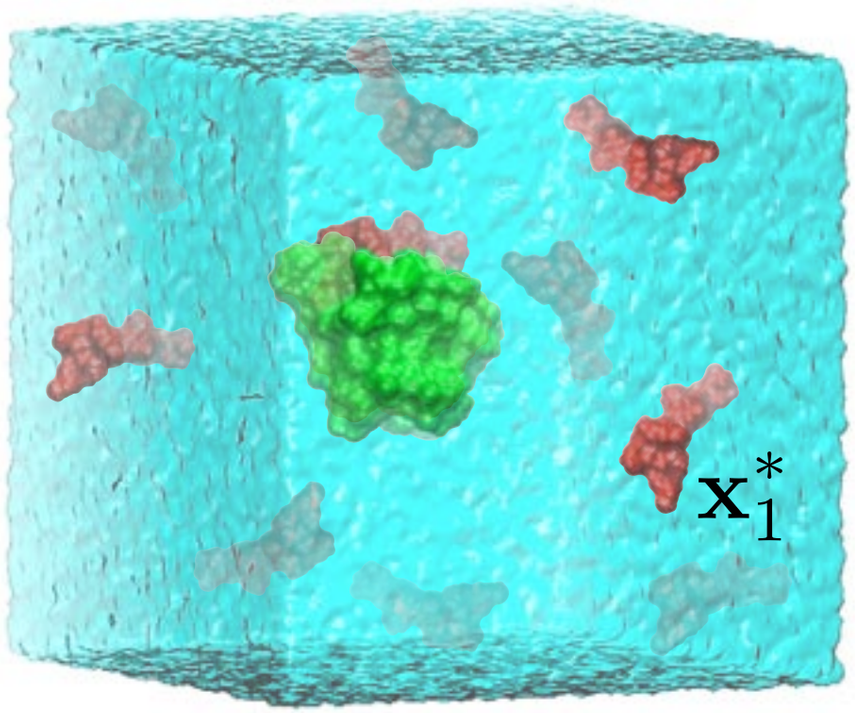
$$[\text{protein}] = p_0\ [\text{protein}]_{\text{tot}}$$

$$[\text{protein} : \text{ligand}] = p_1\ [\text{protein}]_{\text{tot}}$$

Shoup, D.; Szabo, A. *Biophys. J.* **1982**, *40*, 33-39

Woo, H. J.; Roux, B. *Proc. Natl. Acad. Sci. USA* **2005**, 102, 6825-6830

THE LONG-STANDING PROTEIN-LIGAND PROBLEM



$$[\text{ligand}] = N/V_{\text{bulk}}$$

$$K_{\text{eq}} = \frac{1}{[\text{ligand}]} \frac{N \int_{\text{site}} \mathrm{d}\mathbf{1} \int_{\text{bulk}} \mathrm{d}\mathbf{2} \ldots \int_{\text{bulk}} \mathrm{d}\mathbf{N} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta U}}{\int_{\text{bulk}} \mathrm{d}\mathbf{1} \int_{\text{bulk}} \mathrm{d}\mathbf{2} \ldots \int_{\text{bulk}} \mathrm{d}\mathbf{N} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta U}}$$

$$= \frac{1}{[\text{ligand}]} \frac{N \int_{\text{site}} \mathrm{d}\mathbf{1} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta U}}{\int_{\text{bulk}} \mathrm{d}\mathbf{1} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta U}}$$

$$= \frac{1}{[\text{ligand}]} \frac{N \int_{\text{site}} \mathrm{d}\mathbf{1} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta U}}{V_{\text{bulk}} \int_{\text{bulk}} \mathrm{d}\mathbf{1}\ \delta(\mathbf{x}_1 - \mathbf{x}_1^*) \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta U}}$$

$$K_{\text{eq}} = \frac{\int_{\text{site}} \mathrm{d}\mathbf{1} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta U}}{\int_{\text{bulk}} \mathrm{d}\mathbf{1}\ \delta(\mathbf{x}_1 - \mathbf{x}_1^*) \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta U}}$$

alchemical route

geometrical route

Shoup, D.; Szabo, A. *Biophys. J.* **1982**, *40*, 33-39

Woo, H. J.; Roux, B. *Proc. Natl. Acad. Sci. USA* **2005**, 102, 6825-6830

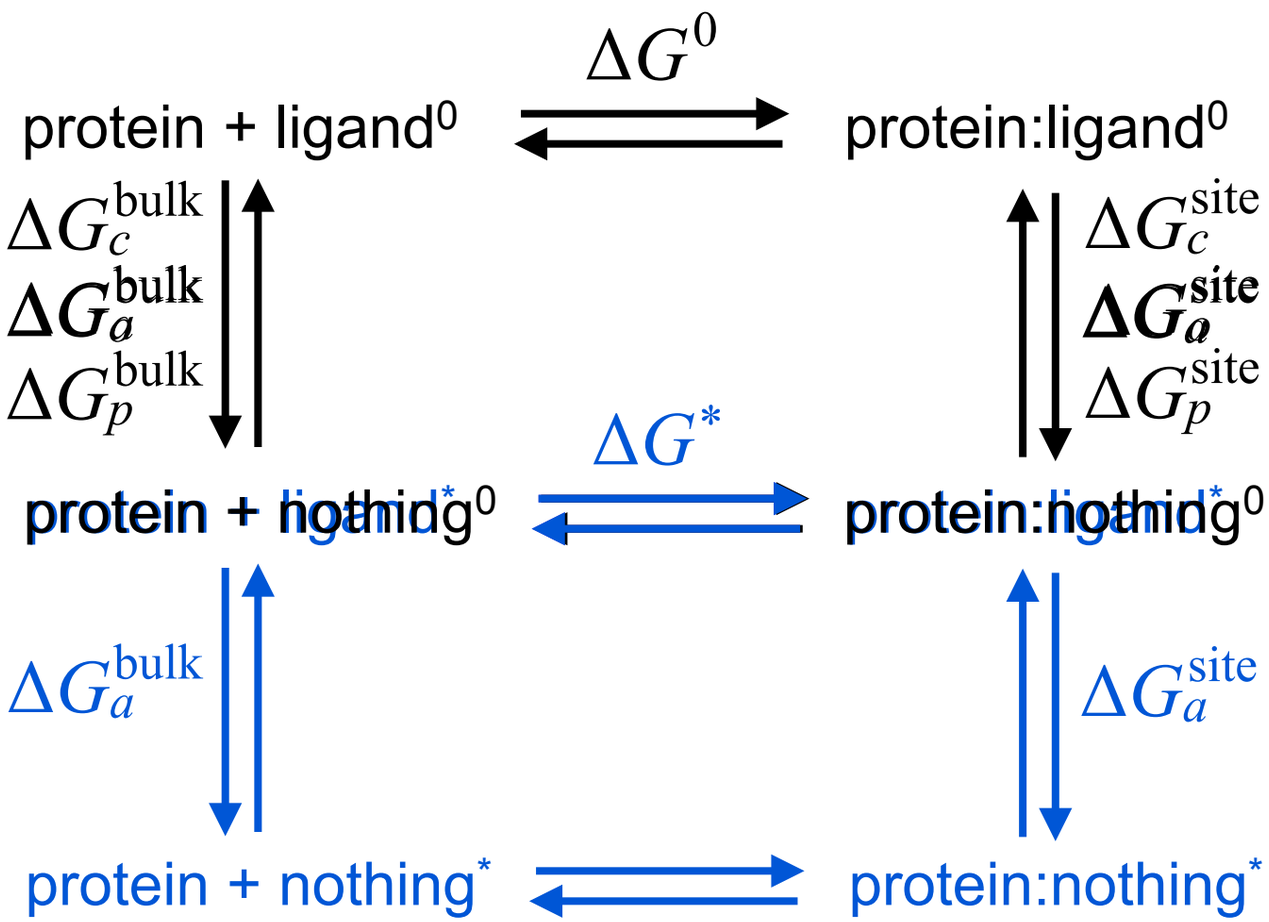
THE LONG-STANDING PROTEIN-LIGAND PROBLEM

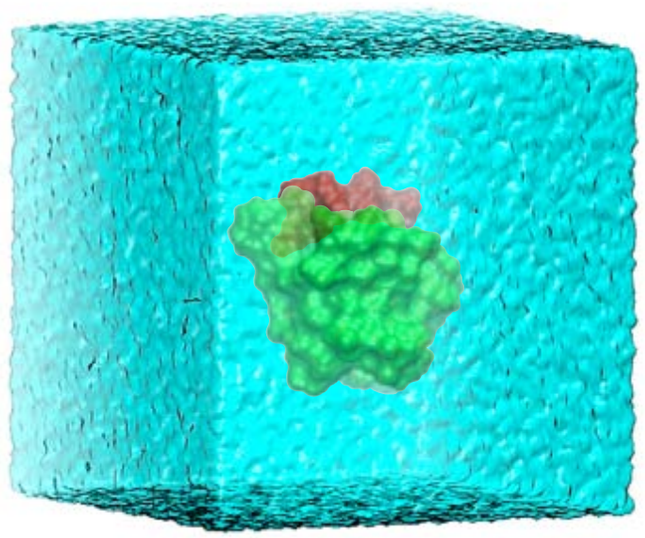
Couple reversibly the ligand to the binding site of the protein

- Floating ligand problem.
- *Corpora non agunt nisi fixata.*

Paul Ehrlich

- Definition of a set of restraints.
- The loss of translational, orientational and conformational entropies contributes to the free energy.

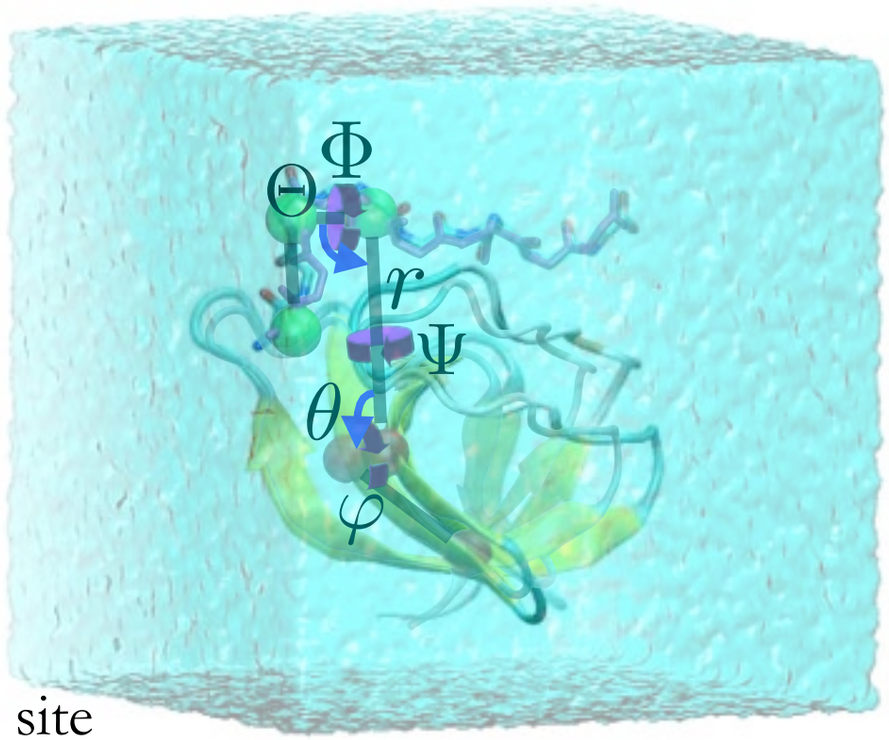




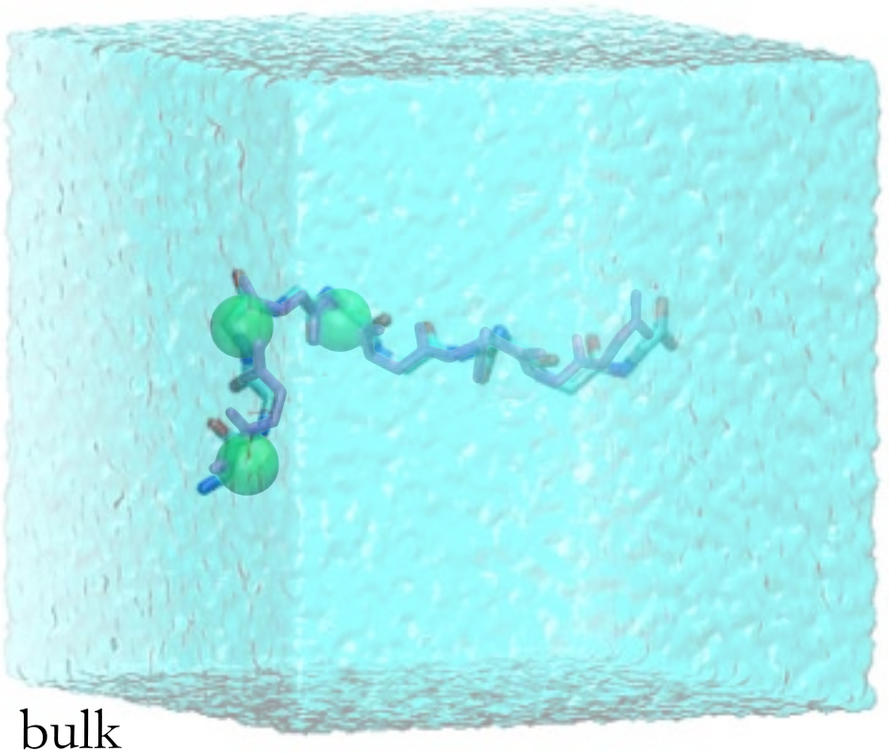
Gilson, M. K. et al. *Biophys. J.*, **1997**, *72*, 1047-1069

Hermans, J.; Wang, L. *J. Am. Chem. Soc.* **1997**, *119*, 2707-2714

THE LONG-STANDING PROTEIN-LIGAND PROBLEM



site



bulk

$$K_{\text{eq}} = \frac{\int_{\text{site}} \mathrm{d}\mathbf{1} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta U_1}}{\int_{\text{site}} \mathrm{d}\mathbf{1} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U_1+u_c)}}$$

$$\times \frac{\int_{\text{site}} \mathrm{d}\mathbf{1} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U_1+u_c)}}{\int_{\text{site}} \mathrm{d}\mathbf{1} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U_1+u_c+u_o)}}$$

$$\times \frac{\int_{\text{site}} \mathrm{d}\mathbf{1} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U_1+u_c+u_o)}}{\int_{\text{site}} \mathrm{d}\mathbf{1} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U_1+u_c+u_o+u_p)}}$$

$$\times \frac{\int_{\text{site}} \mathrm{d}\mathbf{1} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U_1+u_c+u_o+u_p)}}{\int_{\text{site}} \mathrm{d}\mathbf{1} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U_1+u_c+u_o+u_p+u_r)}}$$

$$\times \frac{\int_{\text{site}} \mathrm{d}\mathbf{1} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U_1+u_c+u_o+u_p+u_r)}}{\int_{\text{site}} \mathrm{d}\mathbf{1} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U_0+u_c+u_o+u_p+u_r)}}$$

$$\times \frac{\int_{\text{bulk}} \mathrm{d}\mathbf{1} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U_0+u_c+u_o+u_p+u_r)}}{\int_{\text{bulk}} \mathrm{d}\mathbf{1}\ \delta(\mathbf{x}_1-\mathbf{x}_1^*) \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U_0+u_c+u_o)}}$$

$$\times \frac{\int_{\text{bulk}} \mathrm{d}\mathbf{1}\ \delta(\mathbf{x}_1-\mathbf{x}_1^*) \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U_0+u_c+u_o)}}{\int_{\text{bulk}} \mathrm{d}\mathbf{1}\ \delta(\mathbf{x}_1-\mathbf{x}_1^*) \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U_0+u_c)}}$$

$$\times \frac{\int_{\text{bulk}} \mathrm{d}\mathbf{1}\ \delta(\mathbf{x}_1-\mathbf{x}_1^*) \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U_0+u_c)}}{\int_{\text{bulk}} \mathrm{d}\mathbf{1}\ \delta(\mathbf{x}_1-\mathbf{x}_1^*) \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U_1+u_c)}}$$

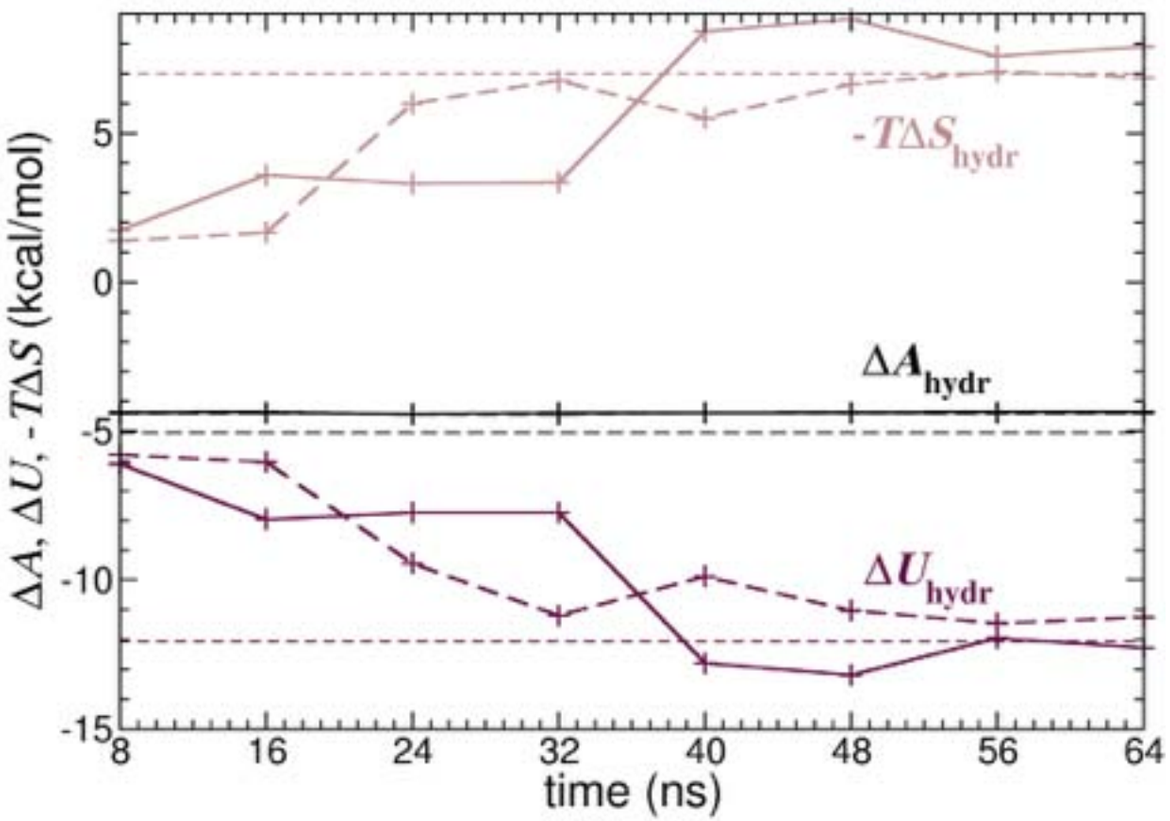
$$\times \frac{\int_{\text{bulk}} \mathrm{d}\mathbf{1}\ \delta(\mathbf{x}_1-\mathbf{x}_1^*) \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U_1+u_c)}}{\int_{\text{bulk}} \mathrm{d}\mathbf{1}\ \delta(\mathbf{x}_1-\mathbf{x}_1^*) \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta U_1}}$$

Deng, Y.; Roux, B. *J. Phys. Chem. B* **2009**, *113*, 2234-2246

HOW ABOUT THE ENTROPY ?

Appreciably more challenging to estimate on account of averages over U_0 and U_1.

$$\Delta S = \frac{1}{T}\left(\frac{\langle U_1 \exp(-\beta\Delta U)\rangle_0}{\langle \exp(-\beta\Delta U)\rangle_0} - \langle U_0\rangle_0\right) + k_B \ln\langle \exp(-\beta\Delta U)\rangle_0$$



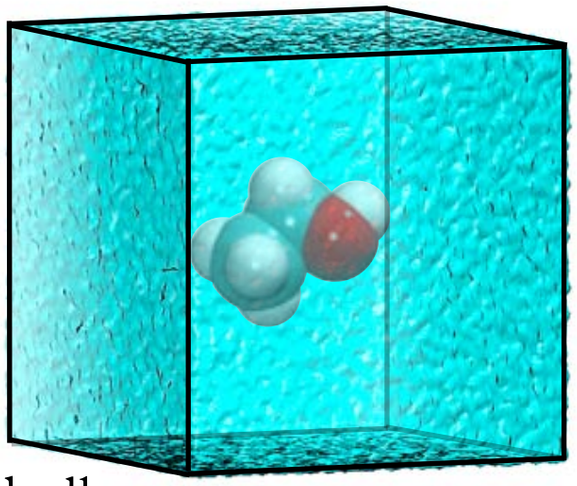
Alternate route:

$$\Delta S = -\left(\frac{\partial \Delta A}{\partial T}\right)_{N,V}$$

Wan, S.; Stote, R. H.; Karplus, M. *J. Chem. Phys.* **2004**, *121*, 9539–9548

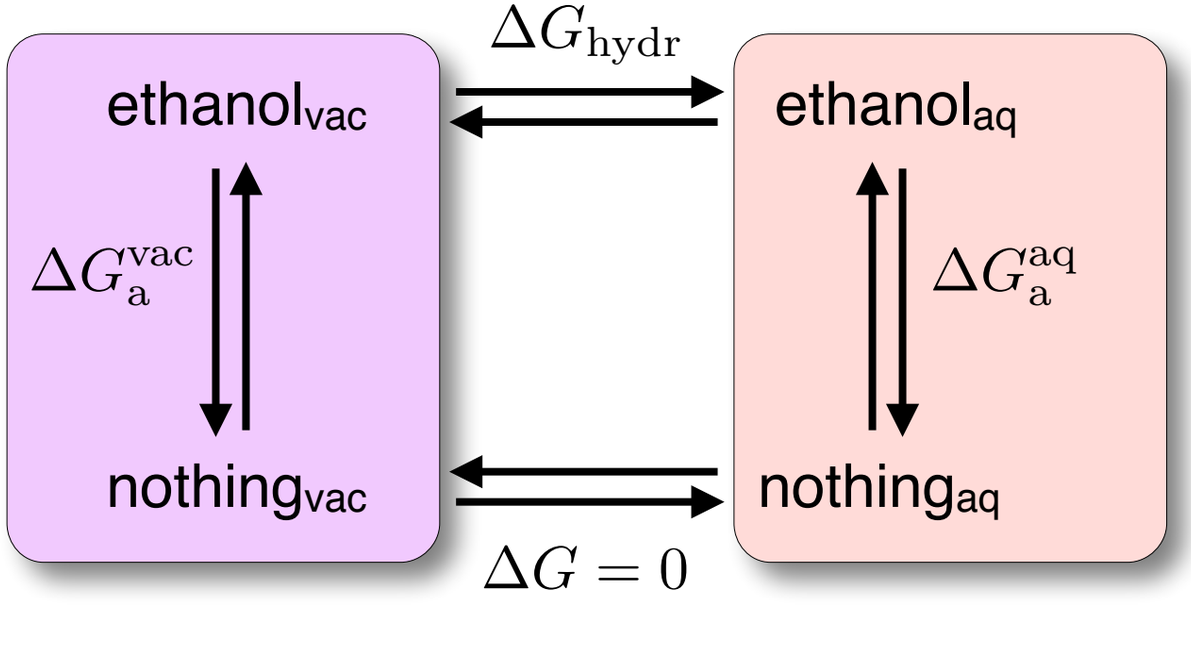
Kubo, M. M.; Gallicchio, E.; Levy, R. M. *J. Phys. Chem. B* **1997**, *101*, 10527-10534

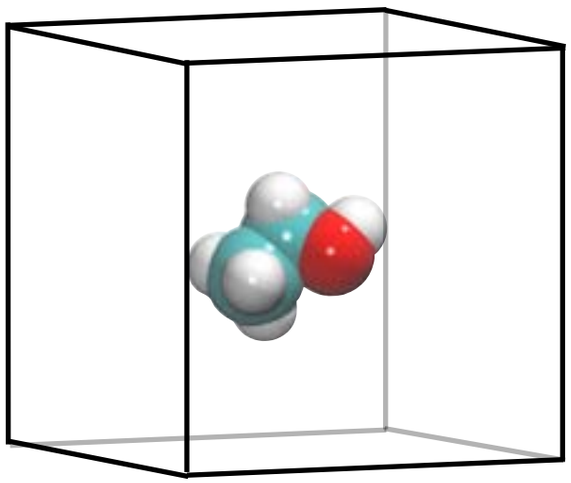
GOOD PRACTICES, GUIDELINES AND RECOMMENDATIONS



bulk

Ethanol hydration





vacuum

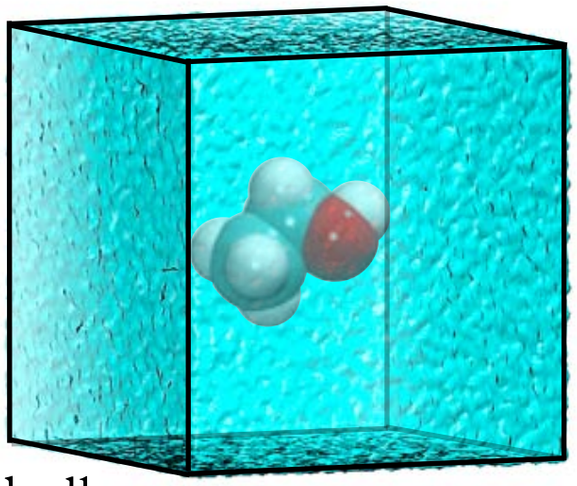
Why do I need to complete the full thermodynamic cycle ?

In different dielectric environments, molecules may adopt very different conformations, corresponding to distinct intramolecular interactions.

In NAMD lingo: `AlchDecouple off`

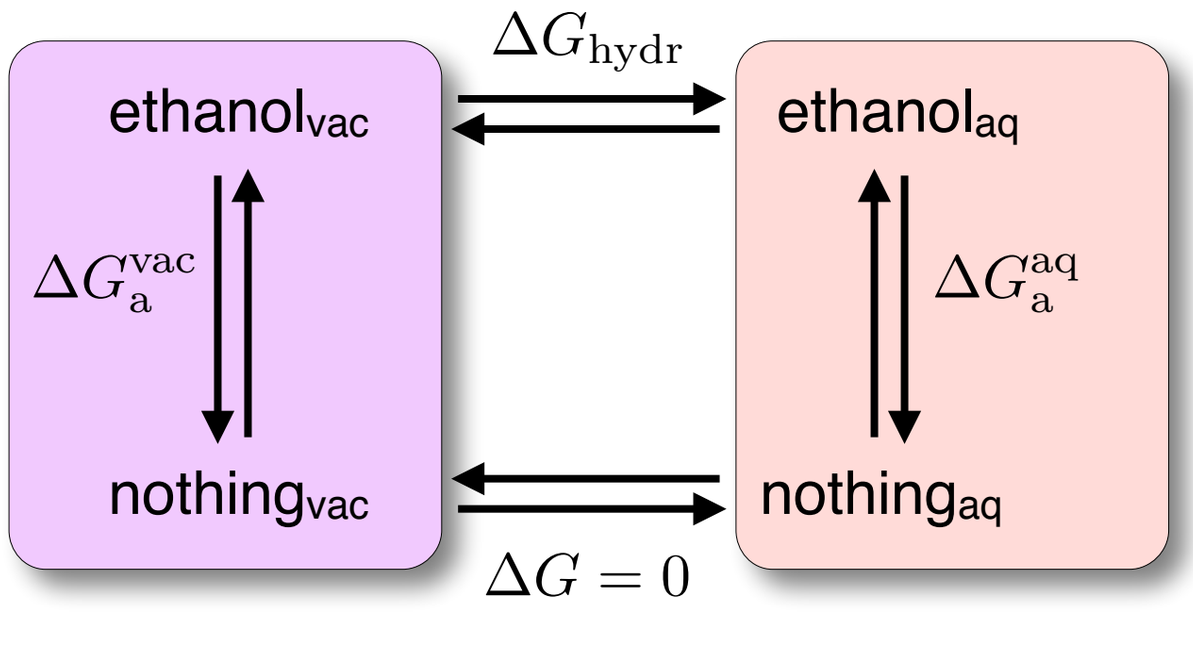
Not a free-energy calculation in vacuum per se, but in a periodic cell bereft of solvent.

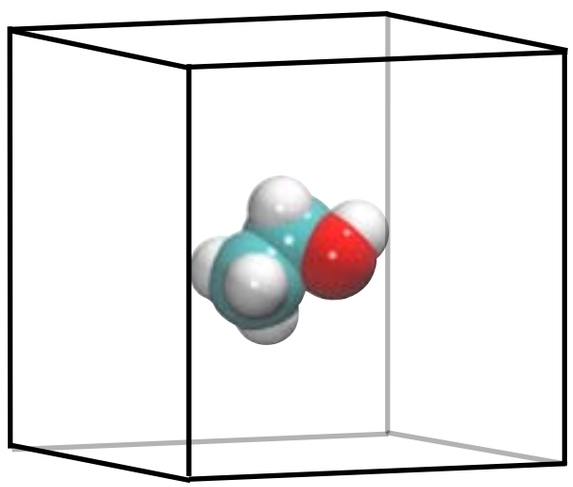
GOOD PRACTICES, GUIDELINES AND RECOMMENDATIONS



bulk

Exercise 2. Ethanol hydration





vacuum

	ΔG (kcal/mol)		
	annihilation	creation	BAR
vacuum	+5.1	-5.3	-5.2
water	+9.5	-9.6	-9.6
hydration	+4.4	-4.3	-4.4

experiment: -5.1 kcal/mol

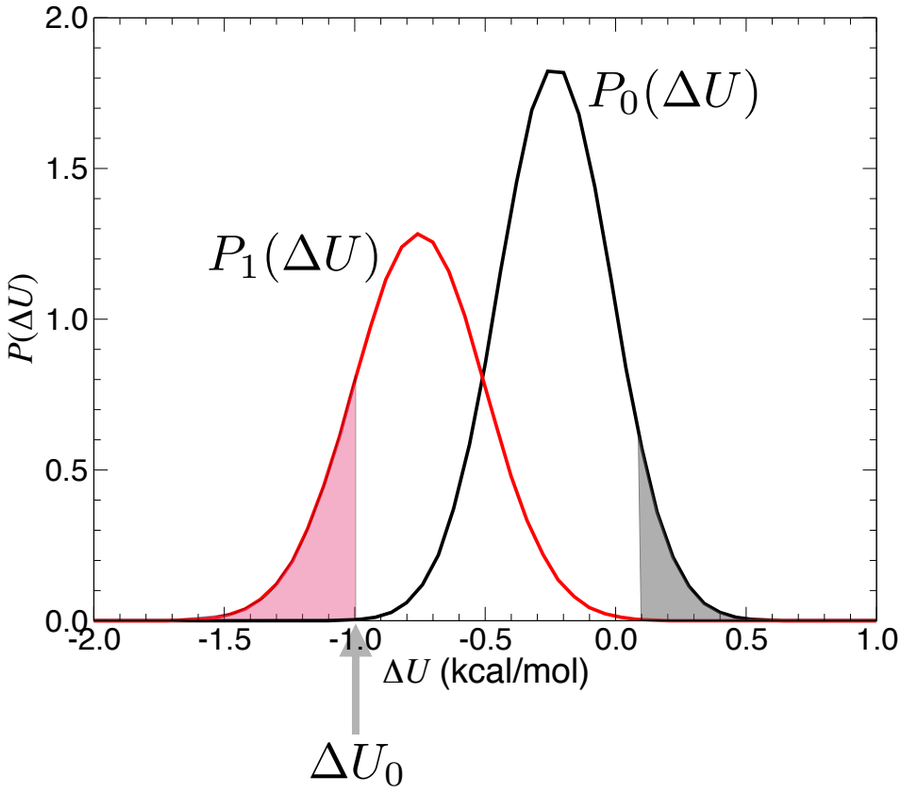
Ben-Naim, A.; Marcus, Y. *J. Chem. Phys.* **1984**, *81*, 2016-2027

GOOD PRACTICES, GUIDELINES AND RECOMMENDATIONS

All free-energy calculations should be accompanied by an error estimate. A distinction between statistical and systematic error ought to be made.

Since the reliability of free-energy estimates depends on the overlap between $P_0(\Delta U)$ and $P_1(\Delta U)$, these distributions should be monitored to assess the degree of overlap.

$$\frac{\delta\epsilon_{\Delta A}}{\exp(-\beta\Delta A)} = -\int_{-\infty}^{\Delta U_0} \mathrm{d}\Delta U \; P_1(\Delta U)$$



Stratification provides an effective, general method for reducing the variance and improving overlap at each stage.

Combining forward and backward simulations using, for instance, the BAR estimator is strongly recommended.

Kofke, D.; Cummings, P. *Fluid Phase Equil.* **1998**, *150*, 41-49

Chipot, C.; Pohorille, A. Free energy calculations. Theory and applications in chemistry and biology, **2007**

Pohorille, A.; Jarzynski, C.; Chipot, C. *J. Phys. Chem. B* **2010**, *114*, 10235-10253

INTRODUCTION

The race for longer and larger simulations
What is the best method for a given problem?

ALCHEMICAL FREE-ENERGY CALCULATIONS

- A tool to address host-guest chemistry questions
- Good practices, guidelines and recommendations
- The long-standing protein-ligand problem

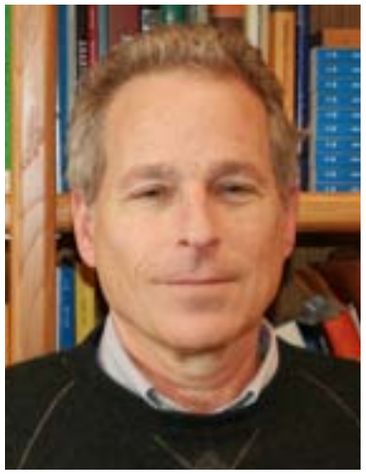
GEOMETRICAL FREE-ENERGY CALCULATIONS

- What is a good reaction-coordinate model?
- A host of methods to measure free-energy changes
- Potentials of mean force and transport phenomena
- Potentials of mean force and recognition and association phenomena
- What about non-equilibrium work computer experiments?

ONGOING CHALLENGES AT THE FRONTIERS OF FREE-ENERGY CALCULATIONS

CONCLUDING REMARKS AND QUESTIONS

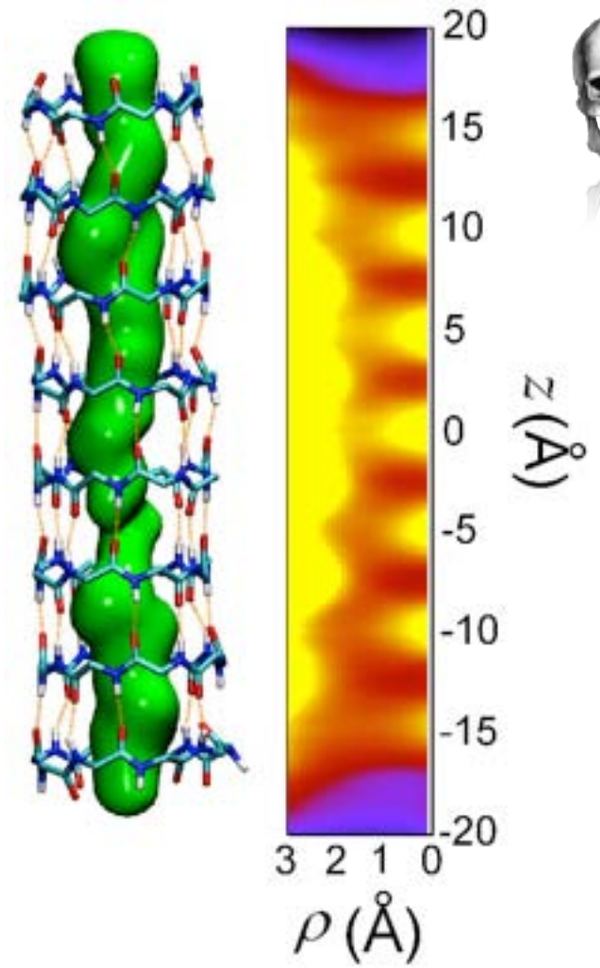
WHAT IS A GOOD REACTION-COORDINATE MODEL ?

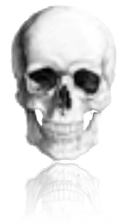


The *true* reaction coordinate generally refers to a unique mathematical object on $\mathbb{R}^{3N}$.

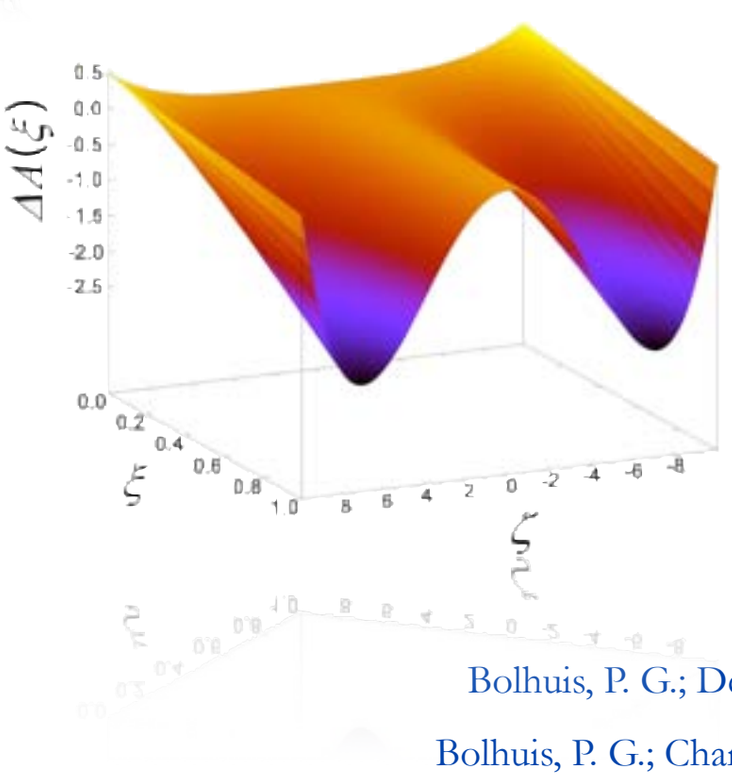
It defines the minimum free-energy pathway connecting the reference state to the target state of the transformation.

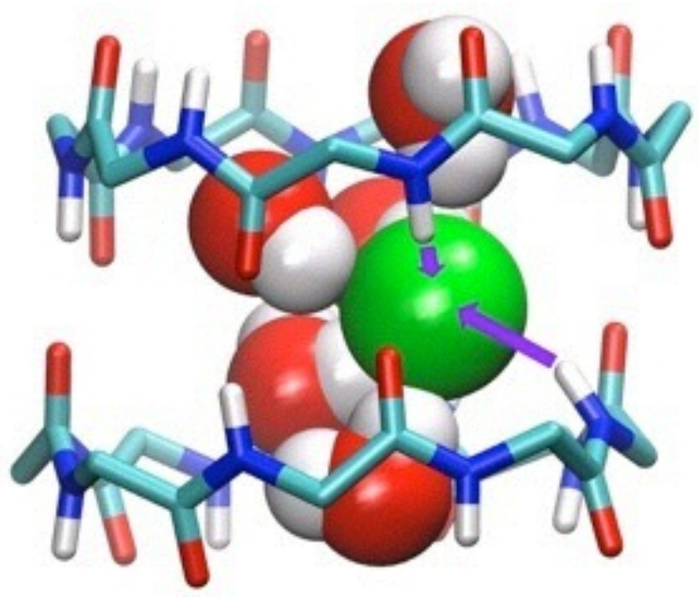
Committor — The probability to reach the target state before returning to the reference state.





A one-dimensional order parameter, namely the long axis of the cavity, is not enough to describe ion conduction in a synthetic channel.

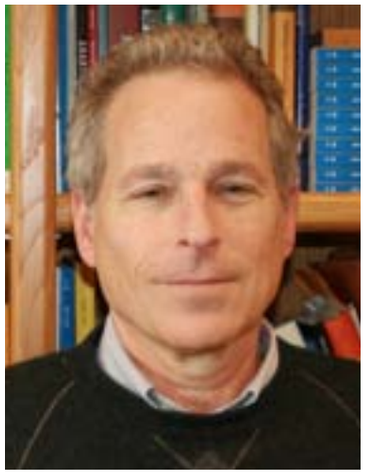




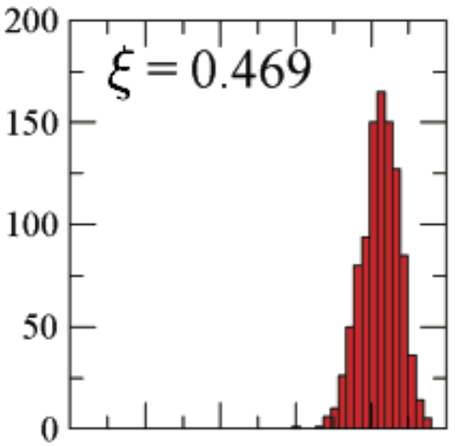
Bolhuis, P. G.; Dellago, C.; Chandler, D. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 5877-5882

Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. *Ann. Rev. Phys. Chem.* **2002**, *59*, 291-318

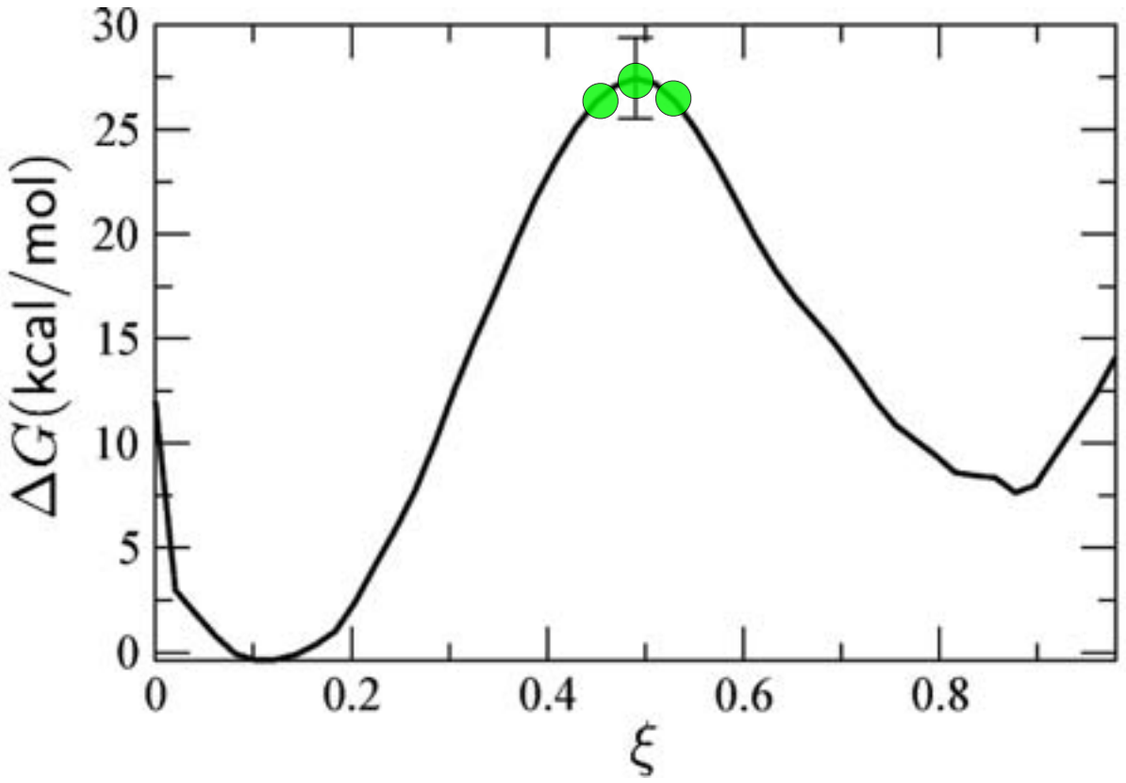
WHAT IS A GOOD REACTION-COORDINATE MODEL ?



While the choice of the reaction-coordinate model does not impact the thermodynamics of the process at hand, it modulates its kinetics.



Option: Determine $N(p_A)$, the distribution of the committor probability, p_A, for the model of the reaction coordinate, ξ.



Run a series of molecular dynamics simulations from the putative maximum of the free-energy barrier and infer $N(p_A)$.

Bolhuis, P. G.; Dellago, C.; Chandler, D. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 5877-5882

Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. *Ann. Rev. Phys. Chem.* **2002**, *59*, 291-318

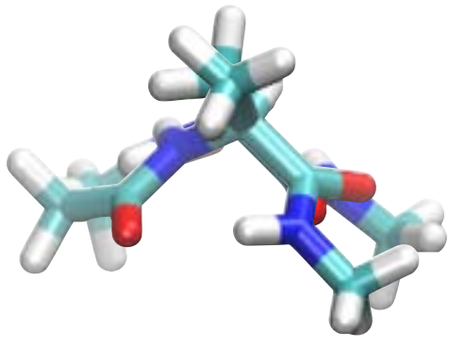
WHAT IS A GOOD REACTION-COORDINATE MODEL ?

low collectivity

`distance`
`distanceZ`
`distanceXY`

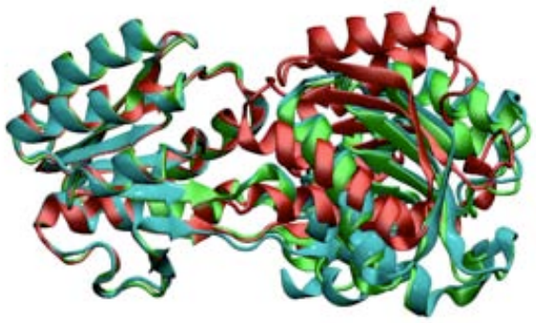
`angle`

`dihedral`



`rmsd`
`gyration`

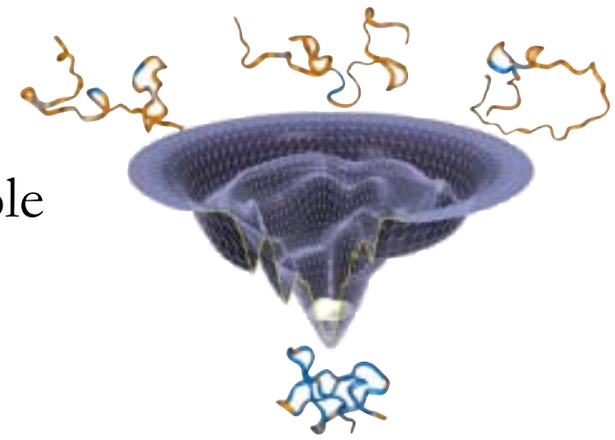
`eigenvector`



high collectivity

Possible linear combination of variables

Degenerate variable



From normal mode or principal component analysis

Hénin, J.; Forin, G.; Chipot, C.; Klein, M. L. *J. Chem. Theor. Comput.* **2010**, *6*, 35-47

Fiorin, G.; Klein, M. L.; Hénin, J. *Mol. Phys.* **2013**, *111*, 3345-3362

A HOST OF METHODS TO MEASURE FREE-ENERGY CHANGES



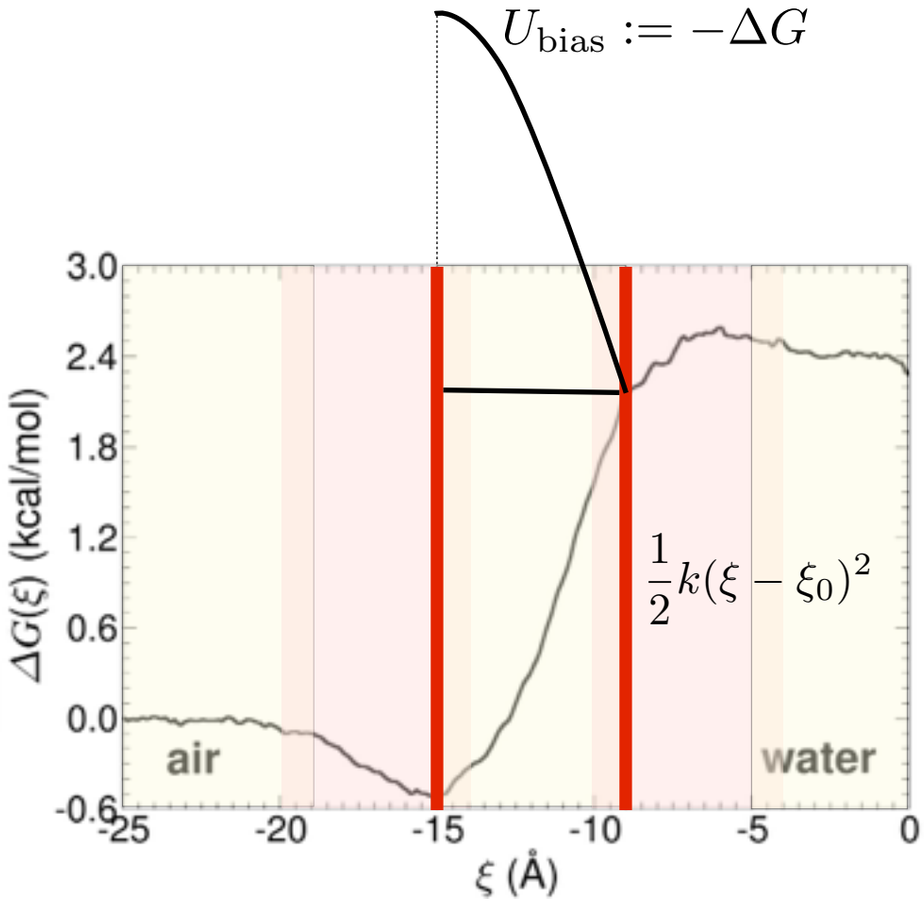
- Conformational flooding.
- Local elevation.
- Metadynamics.


- Umbrella sampling.
- Staging.

$$\Delta G(\xi) = -\frac{1}{\beta}\ln P(\xi) - U_{\text{bias}} + \Delta G_0$$

Weighted histogram analysis method:

$$\begin{cases} P_\lambda(\xi) = \dfrac{\sum_j N_j(\xi)\exp\left(-\beta\sum_i \lambda_i U_{\text{bias},i}\right)}{\sum_k n_k \exp\left(\Delta A_k - \beta\sum_i \lambda_i U_{\text{bias},i}\right)} \\ \exp(-\Delta A_i) = \sum_{U_{\text{bias}}} P_\lambda(\xi) \end{cases}$$



Grubmüller, H. *Phys. Rev. E* **1995**, *52*, 2893-2906
Huber, T. et al. *J. Comput. Aided Mol. Des.* **1994**, *8*, 695-708
Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12562-12565
Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187-199
Valleau, J. P. ; Card, D. N. *J. Chem. Phys.* **1972**, *57*, 5457-5462
Ferrenberg, A. M.; Swendsen, R. H. *Phys. Rev. Lett.* **1989**, *63*, 1195-1198

A HOST OF METHODS TO MEASURE FREE-ENERGY CHANGES



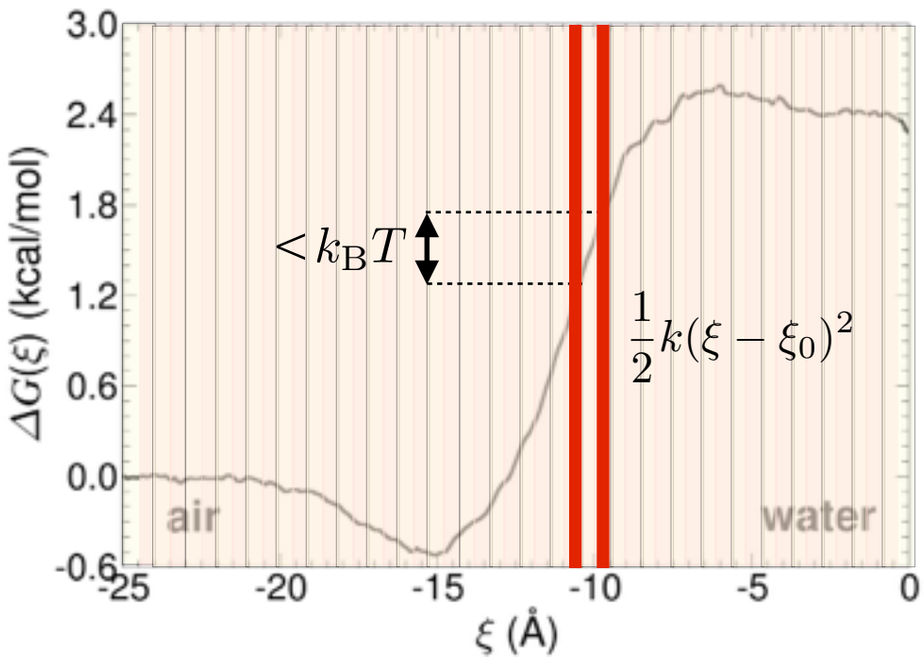
- Conformational flooding.
- Local elevation.
- Metadynamics.


- Umbrella sampling.
- Staging.

$$\Delta G(\xi) = -\frac{1}{\beta}\ln P(\xi) - U_{\text{bias}} + \Delta G_0$$

Weighted histogram analysis method:

$$\begin{cases} P_\lambda(\xi) = \dfrac{\sum_j N_j(\xi)\exp\left(-\beta\sum_i \lambda_i U_{\text{bias},i}\right)}{\sum_k n_k \exp\left(\Delta A_k - \beta\sum_i \lambda_i U_{\text{bias},i}\right)} \\ \exp(-\Delta A_i) = \sum_{U_{\text{bias}}} P_\lambda(\xi) \end{cases}$$



Grubmüller, H. *Phys. Rev. E* **1995**, *52*, 2893-2906

Huber, T. et al. *J. Comput. Aided Mol. Des.* **1994**, *8*, 695-708

Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 12562-12565

Torrie, G. M.; Valleau, J. P. *J. Comput. Phys.* **1977**, *23*, 187-199

Valleau, J. P. ; Card, D. N. *J. Chem. Phys.* **1972**, *57*, 5457-5462

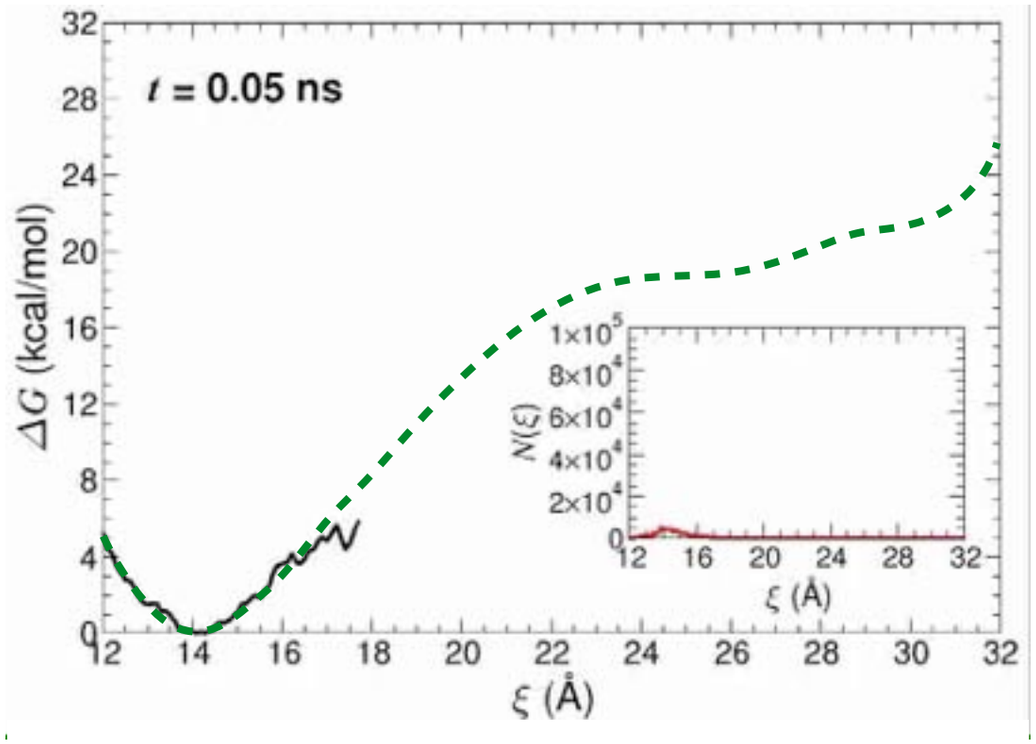
Ferrenberg, A. M.; Swendsen, R. H. *Phys. Rev. Lett.* **1989**, *63*, 1195-1198

A HOST OF METHODS TO MEASURE FREE-ENERGY CHANGES

The derivative of the free energy with respect to the order parameter may be expressed as a sum of configurational averages at constant ξ :

$$\begin{cases} \nabla A(z) := \mathbb{E}\{F(\mathbf{x})|\xi(\mathbf{x}) = z\} \\ F(\mathbf{x}) = \dfrac{\nabla U(\mathbf{x}) \cdot \nabla \xi}{|\nabla \xi|^2} - \dfrac{1}{\beta} \cdot \nabla \left(\dfrac{\nabla \xi}{|\nabla \xi|^2} \right) \end{cases}$$

$\mathbf{v}_{i,i\in[1,\ldots,n]}$ are arbitrarily chosen vector fields of $\mathbb{R}^{3N} \rightarrow \mathbb{R}^{3N}$, which verify $\mathbf{v}_i \cdot \nabla_{\mathbf{x}} \xi_j = \delta_{ij}, \forall i, j$.



The i-th partial derivative of the free energy surface is calculated as the ensemble average of the thermodynamic force:

$$\nabla A(z) = \left\langle \mathbf{v}_i \cdot \nabla U(\mathbf{x}) - \frac{1}{\beta} \nabla \cdot \mathbf{v}_i \right\rangle$$

Carter, E. et al. *Chem. Phys. Lett.* **1989**, *156*, 472-477
Darve, E.; Pohorille, A. *J. Chem. Phys.* **2001**, 115, 9169-9183
Hénin, J.; Chipot, C. *J. Chem. Phys.* **2004**, *121*, 2904-2914
Den Otter, W. *J. Chem. Phys.* **2000**, *112*, 7283-7292

GOOD PRACTICES, GUIDELINES AND RECOMMENDATIONS

Equilibration simulation

Coordinates

`.coor`

Velocities

`.vel`

Simulation cell

`.xsc`

Colvars

`.state`
`.traj`

ABF

`.grad` `.count`
`.pmf`

Colvars

`.in`

Structure

`.psf`

NAMD config

`.namd`

NAMD Scalable Molecular Dynamics

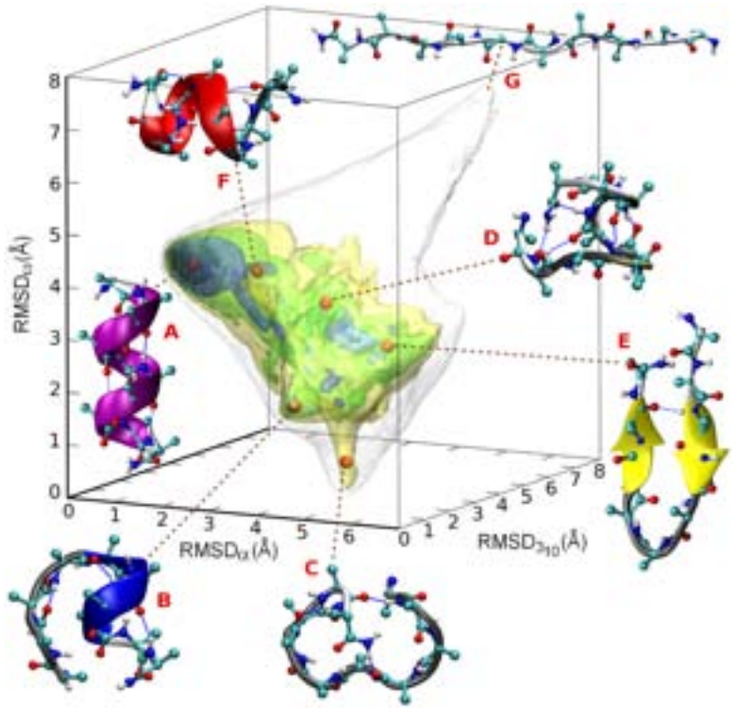
NAMD output

`.log`

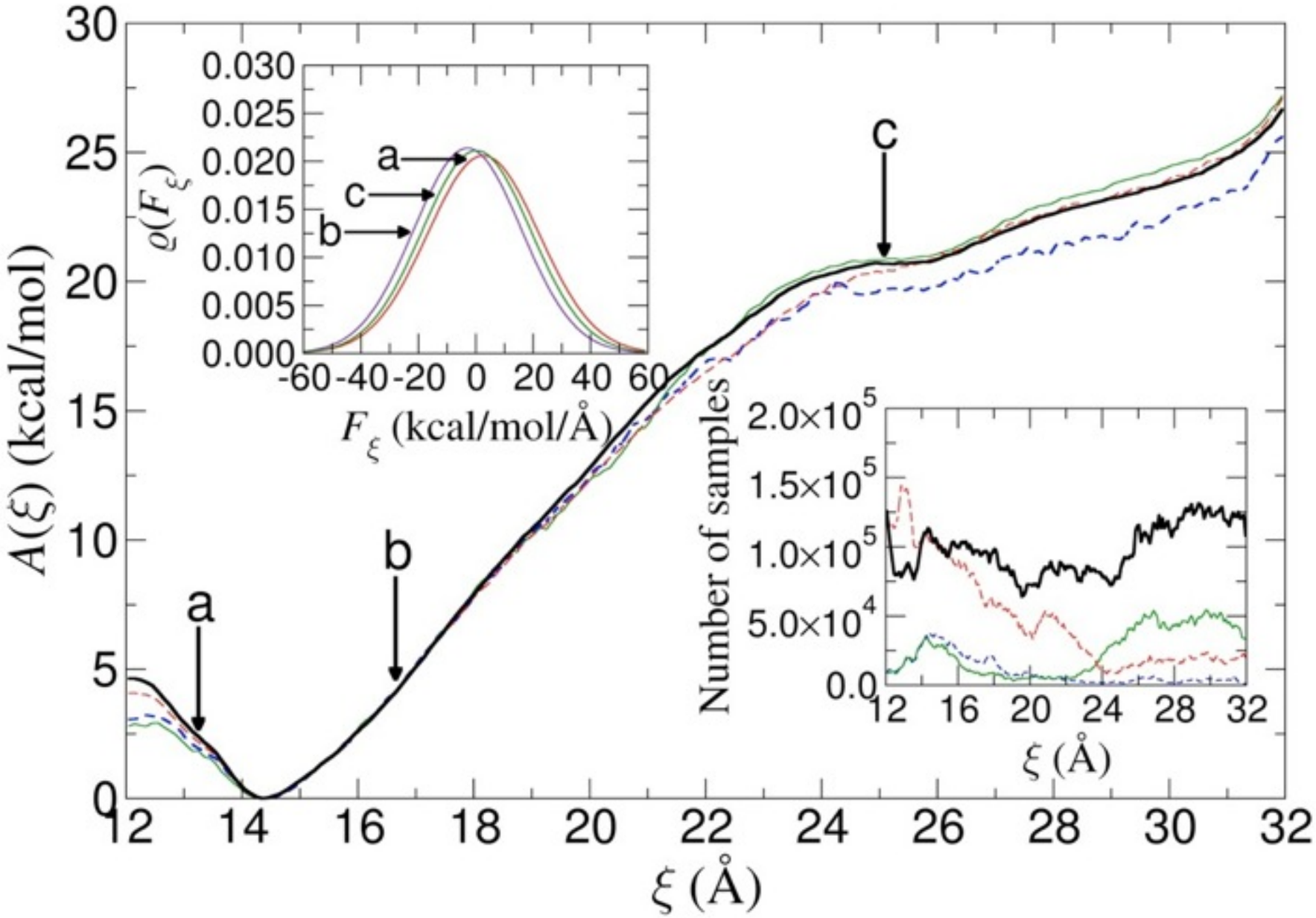
Phillips, J. C. et al. *J. Comput. Chem.* **2005**, *26*, 1781-1802

Bhandarkar, M. et al. NAMD user's guide, version 2.9, **2012**

Reversible unfolding of decaalanine




```
colvar {
   name EndToEndDistance

   width 0.2

   lowerboundary 12.0
   upperboundary 32.0

   lowerwallconstant 100.0
   upperwallconstant 100.0

   outputSystemForce   yes
   outputAppliedForce  yes

   distance {
      group1 {
         atomnumbers { 10 }
      }
      group2 {
         atomnumbers { 92 }
      }
   }
}
```

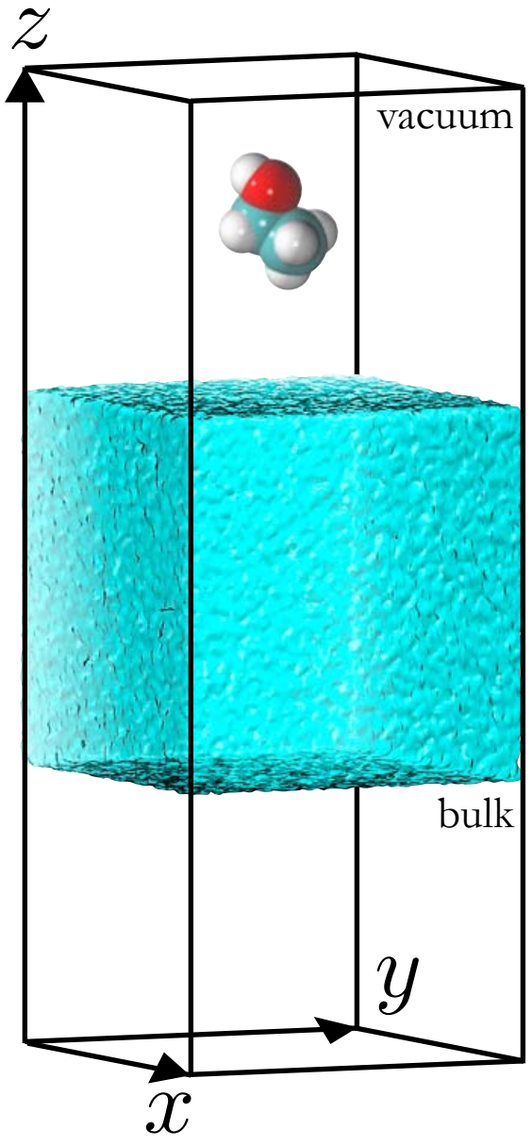


Avoid possible contamination by shaken/rattled degrees of freedom.

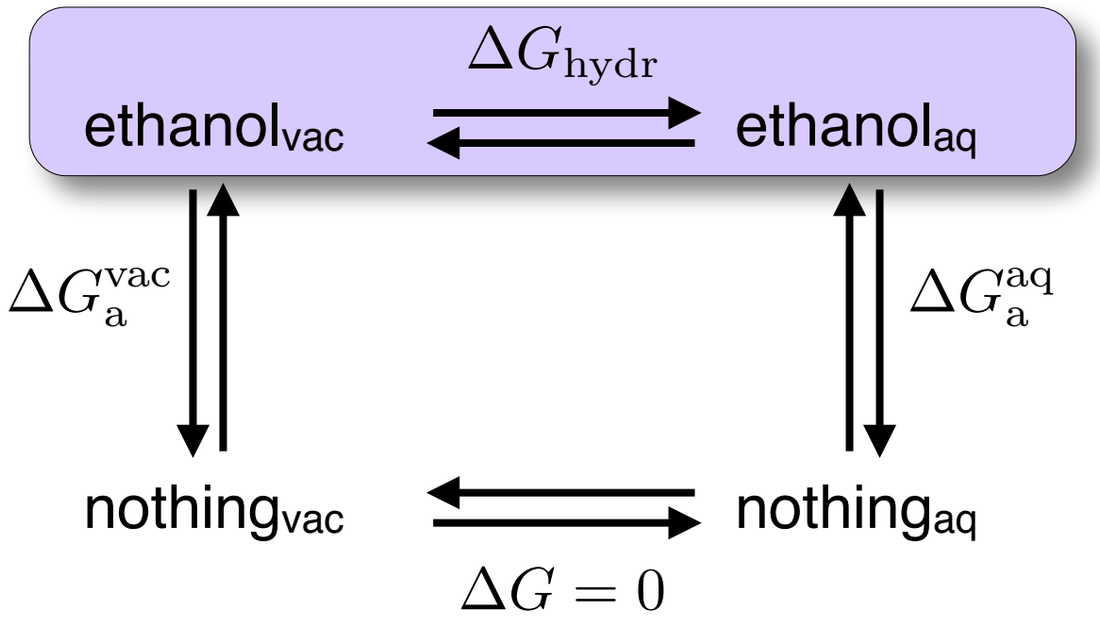
Park, S.; Khalili-Araghi, F.; Tajkhorshid, E.; Schulten, K. *J. Chem. Phys.* **2003**, *119*, 3559-3566

Hénin, J.; Chipot, C. *J. Chem. Phys.* **2004**, *121*, 2904-2914

GOOD PRACTICES, GUIDELINES AND RECOMMENDATIONS



Back to exercise 2. Ethanol hydration



Access to both hydration and adsorption (experiment: +2.5 kcal/mol) free energies.

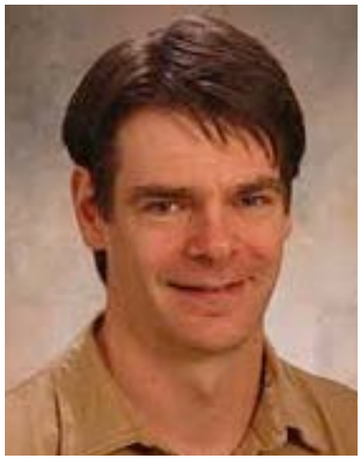
Decomposition of $\nabla_z A$ and projection onto z yields free-energy contributions.

```
colvar {
   name ProjectionZ
   width 0.1
   lowerboundary  0.0
   upperboundary  5.0
   lowerwallconstant 100.0
   upperwallconstant 100.0

   distanceZ {
      ref {
         atomsFile reference.pdb
         atomsCol B
      }
      main  {
         atomnumbers { 1 2 3 4 5 6 7 8 9 }
      }
   }
}

abf {
  colvars ProjectionZ
  fullSamples    1000
}
```


Chipot, C.; Pohorille, A. Free energy calculations. Theory and applications in chemistry and biology, **2007**

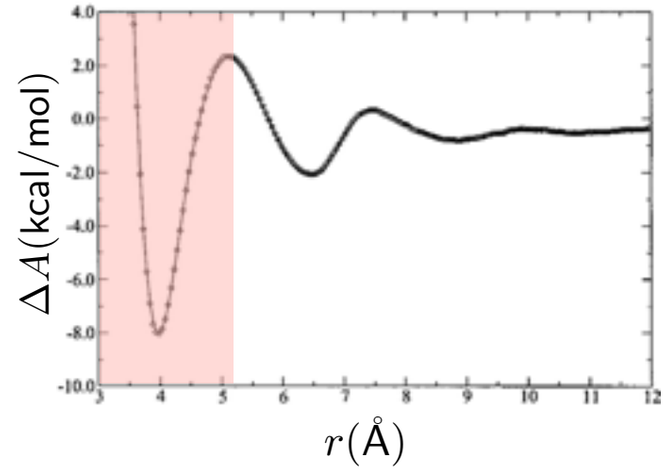
Wilson, M. A.; Pohorille, A. *J. Phys. Chem. B* **1997**, *101*, 3130-3135

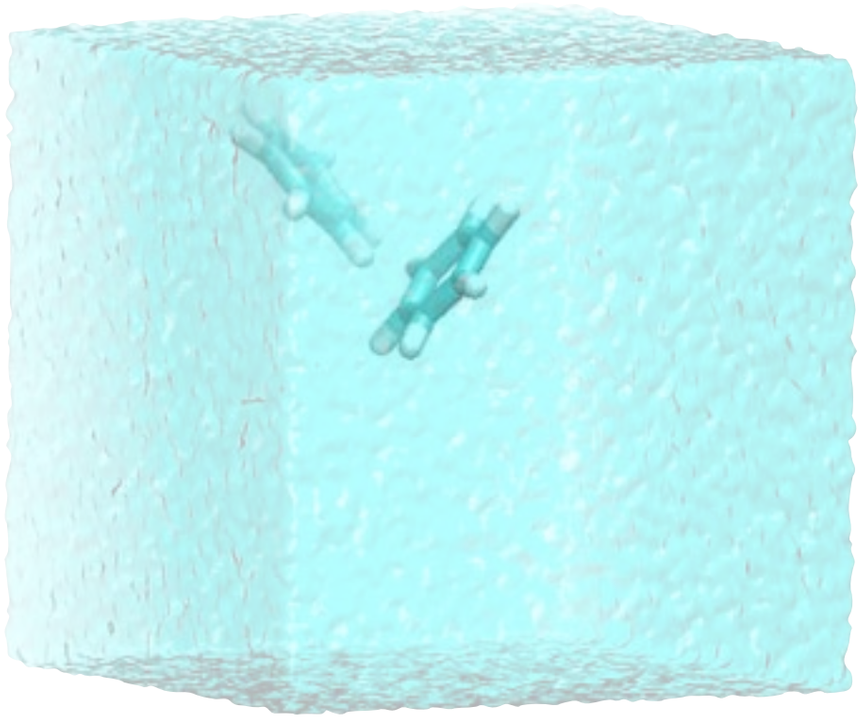
GOOD PRACTICES, GUIDELINES AND RECOMMENDATIONS



Measuring binding constants from one-dimensional separation potentials of mean force is justified in the limit of all other degrees of freedom being appropriately sampled. This is true for small, fast-relaxing molecular species.

$$K_a = 4\pi \int_0^{R_c} \mathrm{d}r \; r^2 \; \exp[-\beta \Delta A(r)]$$

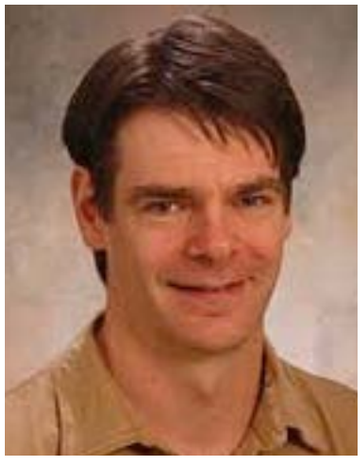




Shoup, D.; Szabo, A. *Biophys. J.* **1982**, *40*, 33-39

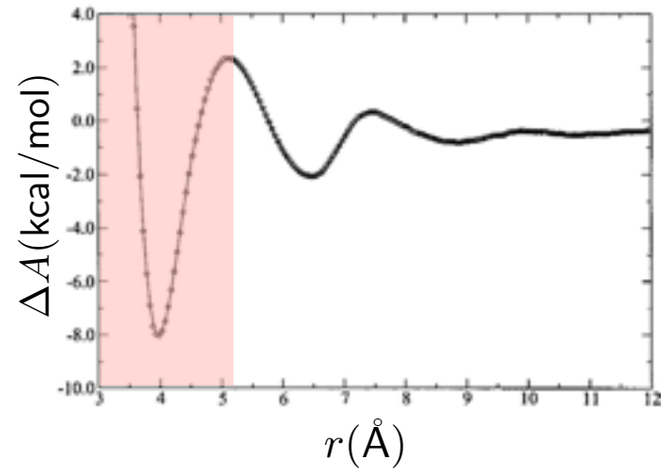
Woo, H. J.; Roux, B. *Proc. Natl. Acad. Sci. USA* **2005**, 102, 6825-6830

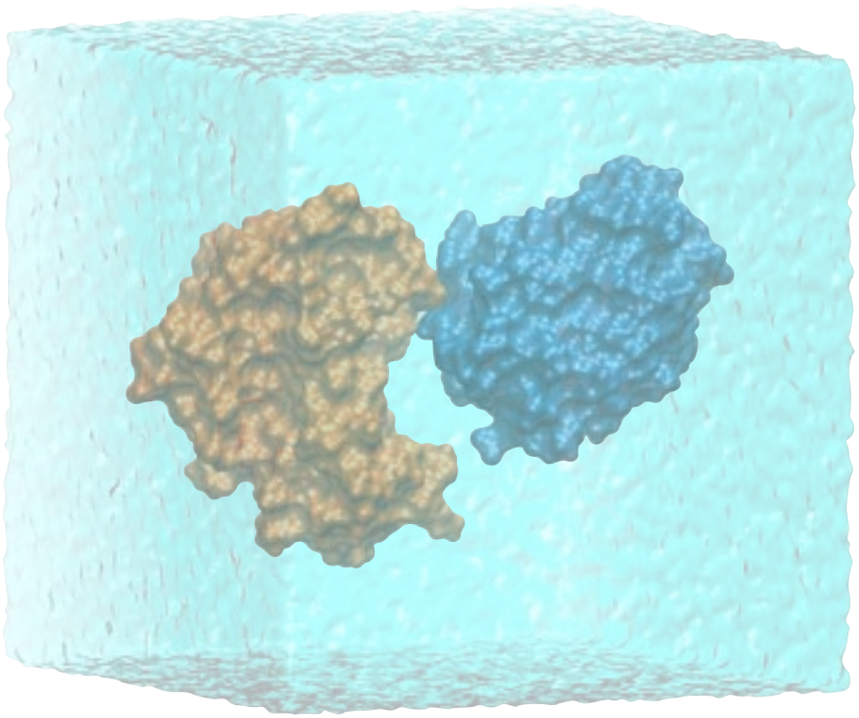
GOOD PRACTICES, GUIDELINES AND RECOMMENDATIONS



Measuring binding constants from one-dimensional separation potentials of mean force is justified in the limit of all other degrees of freedom being appropriately sampled. This is true for small, fast-relaxing molecular species.

$$K_a = 4\pi \int_0^{R_c} \mathrm{d}r \; r^2 \; \exp[-\beta \Delta A(r)]$$

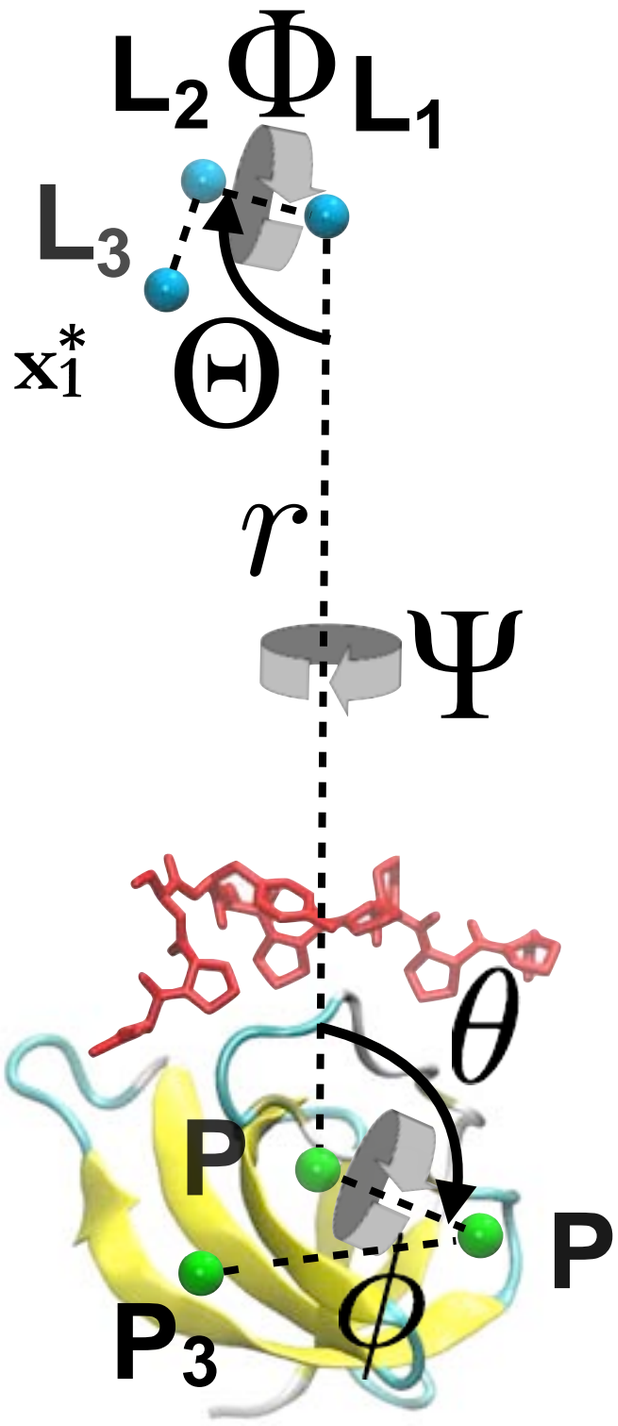




In more complex molecular assemblies, e.g., protein-ligand complexes, the partners acquire upon separation additional configurational - i.e., conformational, positional and orientational entropy, not easily captured over timescales amenable to molecular dynamics.

Shoup, D.; Szabo, A. *Biophys. J.* **1982**, *40*, 33-39

Gumbart, J. C.; Roux, B.; Chipot, C. *J. Chem. Theor. Comput.* **2013**, *9*, 3789-3798



$$K_{\text{eq}} = \frac{\int_{\text{site}} \mathrm{d}\mathbf{1} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta U}}{\int_{\text{site}} \mathrm{d}\mathbf{1} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U+u_c)}}$$

$$\times \frac{\int_{\text{site}} \mathrm{d}\mathbf{1} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U+u_c)}}{\int_{\text{site}} \mathrm{d}\mathbf{1} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U+u_c+u_\Theta)}}$$

$$\times \frac{\int_{\text{site}} \mathrm{d}\mathbf{1} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U+u_c+u_\Theta)}}{\int_{\text{site}} \mathrm{d}\mathbf{1} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U+u_c+u_\Theta+u_\Phi)}}$$

$$\times \frac{\int_{\text{site}} \mathrm{d}\mathbf{1} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U+u_c+u_\Theta+u_\Phi)}}{\int_{\text{site}} \mathrm{d}\mathbf{1} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U+u_c+u_\Theta+u_\Phi+u_\Psi)}}$$

$$\times \frac{\int_{\text{site}} \mathrm{d}\mathbf{1} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U+u_c+u_o)}}{\int_{\text{site}} \mathrm{d}\mathbf{1} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U+u_c+u_o+u_\theta)}}$$

$$\times \frac{\int_{\text{site}} \mathrm{d}\mathbf{1} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U+u_c+u_o+u_\theta)}}{\int_{\text{site}} \mathrm{d}\mathbf{1} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U+u_c+u_o+u_\theta+u_\phi)}}$$

$$\times \frac{\int_{\text{site}} \mathrm{d}\mathbf{1} \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U+u_c+u_o+u_p)}}{\int_{\text{bulk}} \mathrm{d}\mathbf{1}\ \delta(\mathbf{x}_1-\mathbf{x}_1^*) \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U+u_c+u_o)}}$$

$$\times \frac{\int_{\text{bulk}} \mathrm{d}\mathbf{1}\ \delta(\mathbf{x}_1-\mathbf{x}_1^*) \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U+u_c+u_\Theta+u_\Phi+u_\Psi)}}{\int_{\text{bulk}} \mathrm{d}\mathbf{1}\ \delta(\mathbf{x}_1-\mathbf{x}_1^*) \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U+u_c+u_\Theta+u_\Phi)}}$$

$$\times \frac{\int_{\text{bulk}} \mathrm{d}\mathbf{1}\ \delta(\mathbf{x}_1-\mathbf{x}_1^*) \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U+u_c+u_\Theta+u_\Phi)}}{\int_{\text{bulk}} \mathrm{d}\mathbf{1}\ \delta(\mathbf{x}_1-\mathbf{x}_1^*) \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U+u_c+u_\Theta)}}$$

$$\times \frac{\int_{\text{bulk}} \mathrm{d}\mathbf{1}\ \delta(\mathbf{x}_1-\mathbf{x}_1^*) \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U+u_c+u_\Theta)}}{\int_{\text{bulk}} \mathrm{d}\mathbf{1}\ \delta(\mathbf{x}_1-\mathbf{x}_1^*) \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U+u_c)}}$$

$$\times \frac{\int_{\text{bulk}} \mathrm{d}\mathbf{1}\ \delta(\mathbf{x}_1-\mathbf{x}_1^*) \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta(U+u_c)}}{\int_{\text{bulk}} \mathrm{d}\mathbf{1}\ \delta(\mathbf{x}_1-\mathbf{x}_1^*) \int \mathrm{d}\mathbf{x}\ \mathrm{e}^{-\beta U}}$$

Yu, Y. B. et al. *Biophys. J.* **2001**, *81*, 1632-1642

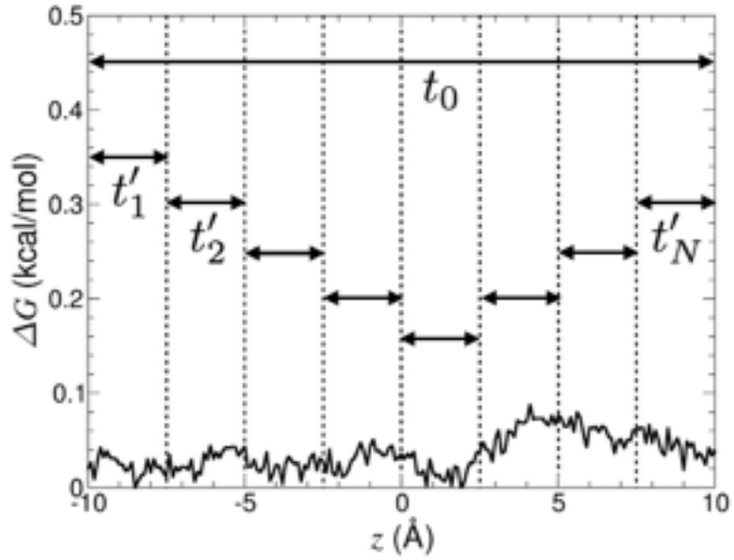
Woo, H. J.; Roux, B. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 6825-6830

Gumbart, J. C.; Roux, B.; Chipot, C. *J. Chem. Theory Comput.* **2013,** *9*, 794-802

GOOD PRACTICES, GUIDELINES AND RECOMMENDATIONS

ξ ought to be completely decoupled from degrees of freedom to which holonomic constraints are applied.

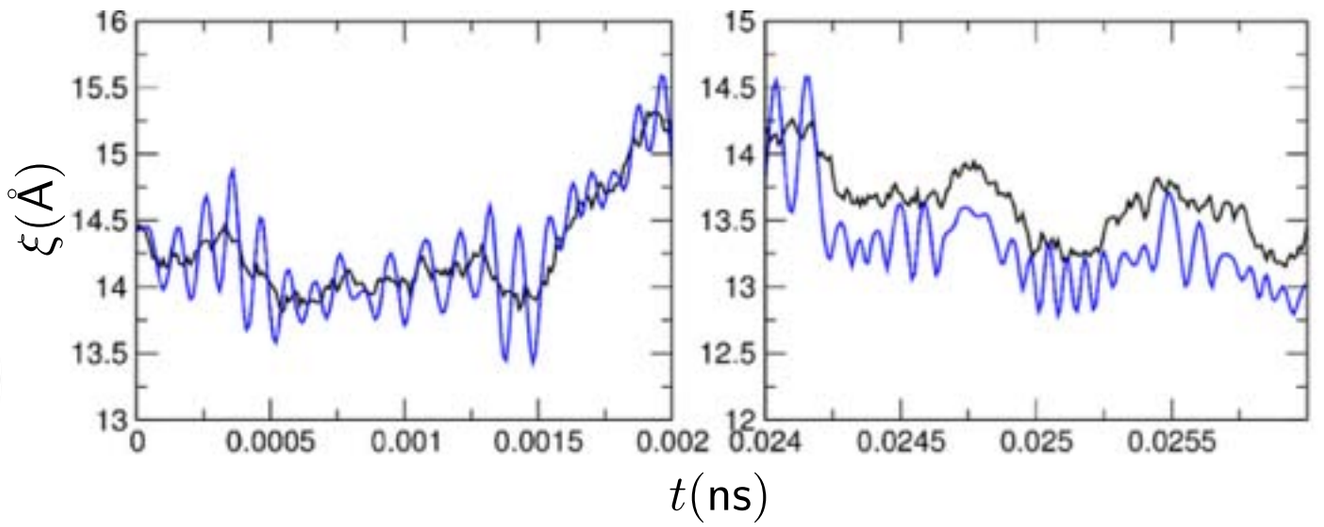
As a matter of principle, ξ ought to be stratified: $t_0 > \sum_i t'_i$



Turn to the extended-Lagrangian formulation of ABF in the event geometrical restraints are coupled to ξ.

```
ExtendedLagrangian    on
```

$$F(\mathbf{x}) = \frac{\nabla U(\mathbf{x}) \cdot \nabla \xi}{|\nabla \xi|^2} - \frac{1}{\beta} \cdot \nabla \left(\frac{\nabla \xi}{|\nabla \xi|^2} \right)$$



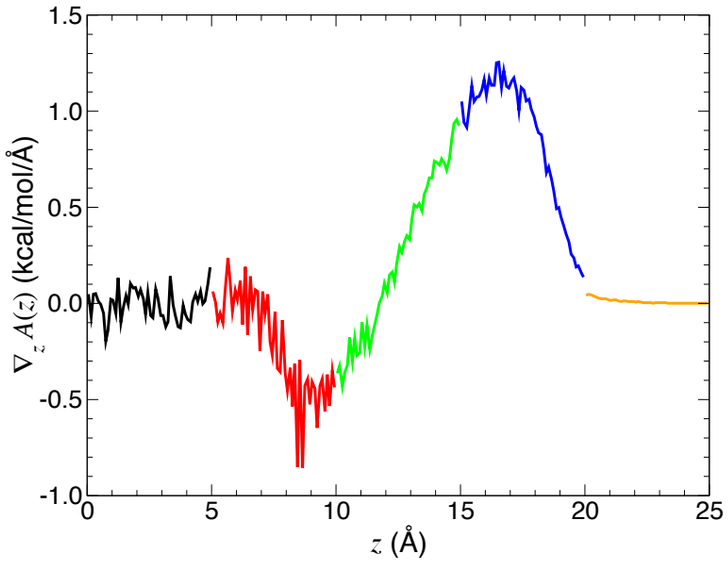
Chipot, C.; Pohorille, A. Free energy calculations. Theory and applications in chemistry and biology, **2007**

Lelièvre, T.; Stoltz, G.; Rousset, M. Free energy computations: A mathematical perspective, **2010**

Comer, J.; Gumbart, J. C.; Hénin, J.; Lelièvre, T.; Pohorille, A.; Chipot, C. *J. Phys. Chem.* **2014**

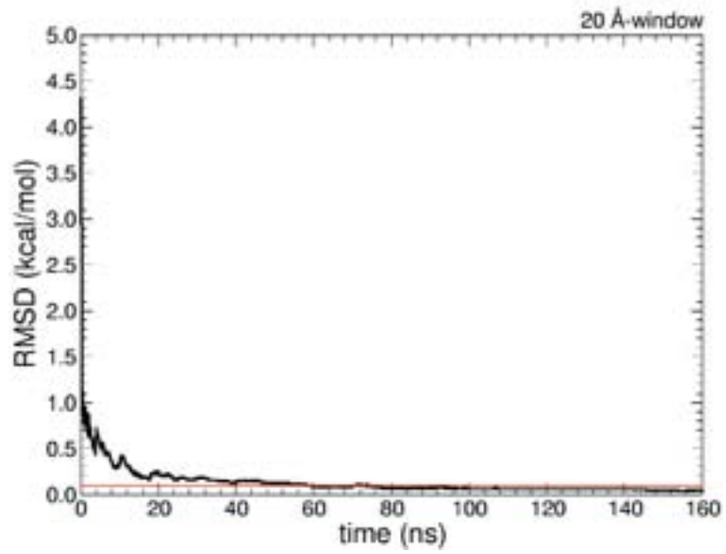
GOOD PRACTICES, GUIDELINES AND RECOMMENDATIONS

In stratified ABF calculations, continuity of the average force ought to be verified.



Free-energy profiles ought to be provided with error bars, $\sigma_{\Delta A} \simeq \frac{\sigma}{N^{1/2}}(1+2\kappa)^{1/2}$

Assess convergence, for instance, by breaking down the free-energy calculations in multiple sub-runs and compute a root mean-square deviation with respect to the last one.



Chipot, C.; Pohorille, A. Free energy calculations. Theory and applications in chemistry and biology, **2007**

Lelièvre, T.; Stoltz, G.; Rousset, M. Free energy computations: A mathematical perspective, **2010**

Comer, J.; Gumbart, J. C.; Hénin, J.; Lelièvre, T.; Pohorille, A.; Chipot, C. *J. Phys. Chem.* **2014**

WHAT ABOUT NON-EQUILIBRIUM WORK COMPUTER EXPERIMENTS ?

$$\exp(-\beta\Delta A) = \langle \exp(-\beta w) \rangle$$

Pulling simulations are usually carried out at a velocity about 10^3 greater than that of experiment.

$$\frac{P_0(+w)}{P_1(-w)} = \exp[+\beta(w - \Delta A)]$$

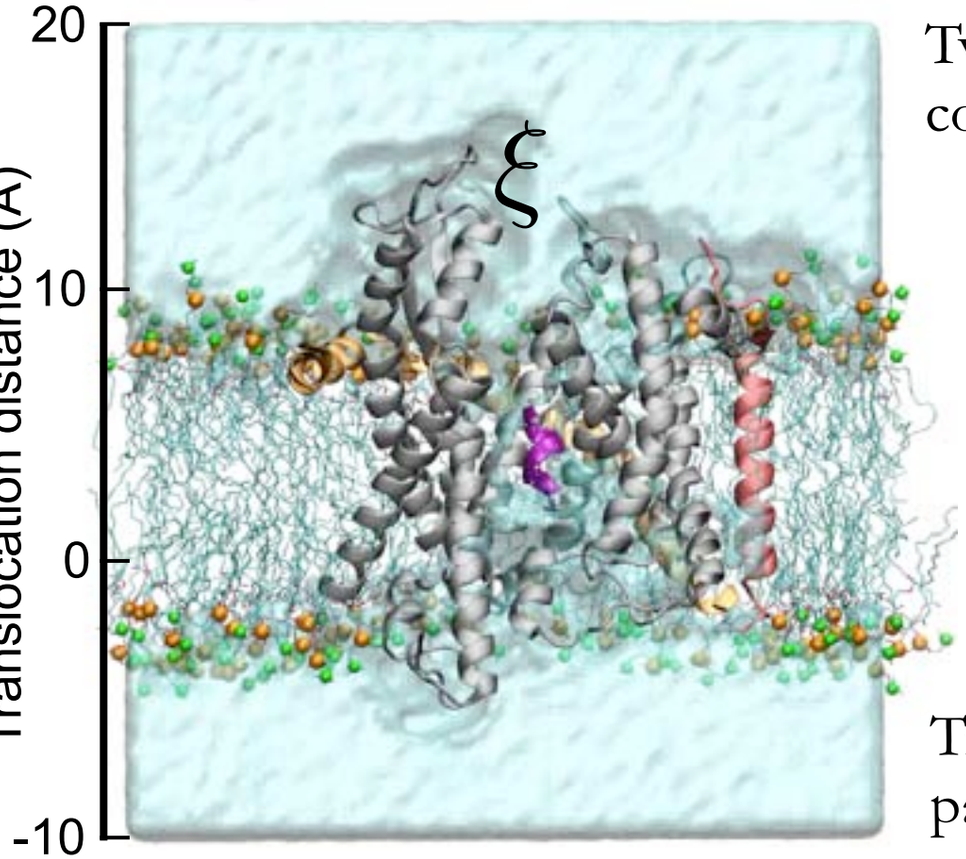
Akin to BAR, combine the forward and backward transformations to reduce the variance.

Though cumulant expansions help, the non-equilibium work route to free-energy differences requires near-equilibrium conditions to converge.

In general, there is no fundamental reason to resort to non-equilibrium work experiments if the free-energy change can be estimated at equilibrium.

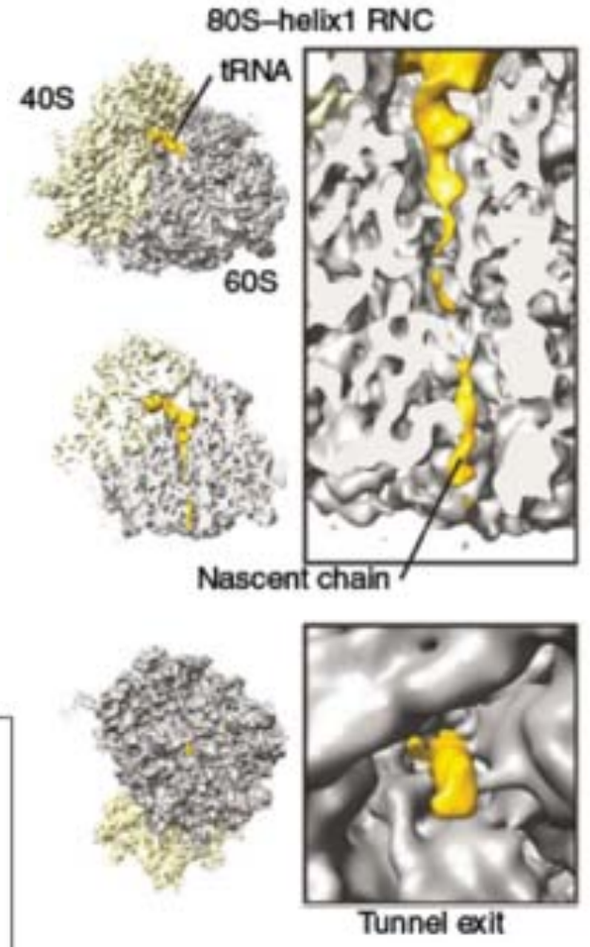
Jarzynski, C. *Phys. Rev. Lett.* **1997**, *78*, 2690-2693

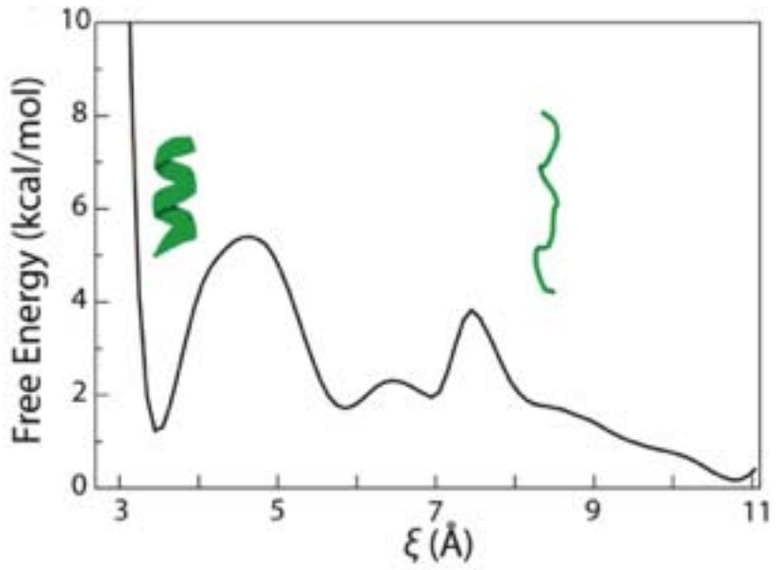
Crooks, G. *J. Stat. Phys.* **1998**, *90*, 1481-1487



Two-dimensional reaction coordinate: Translation in SecY plus the concerted folding of the peptide chain (ξ).

Translocation of proteins supposes partial opening of SecY.





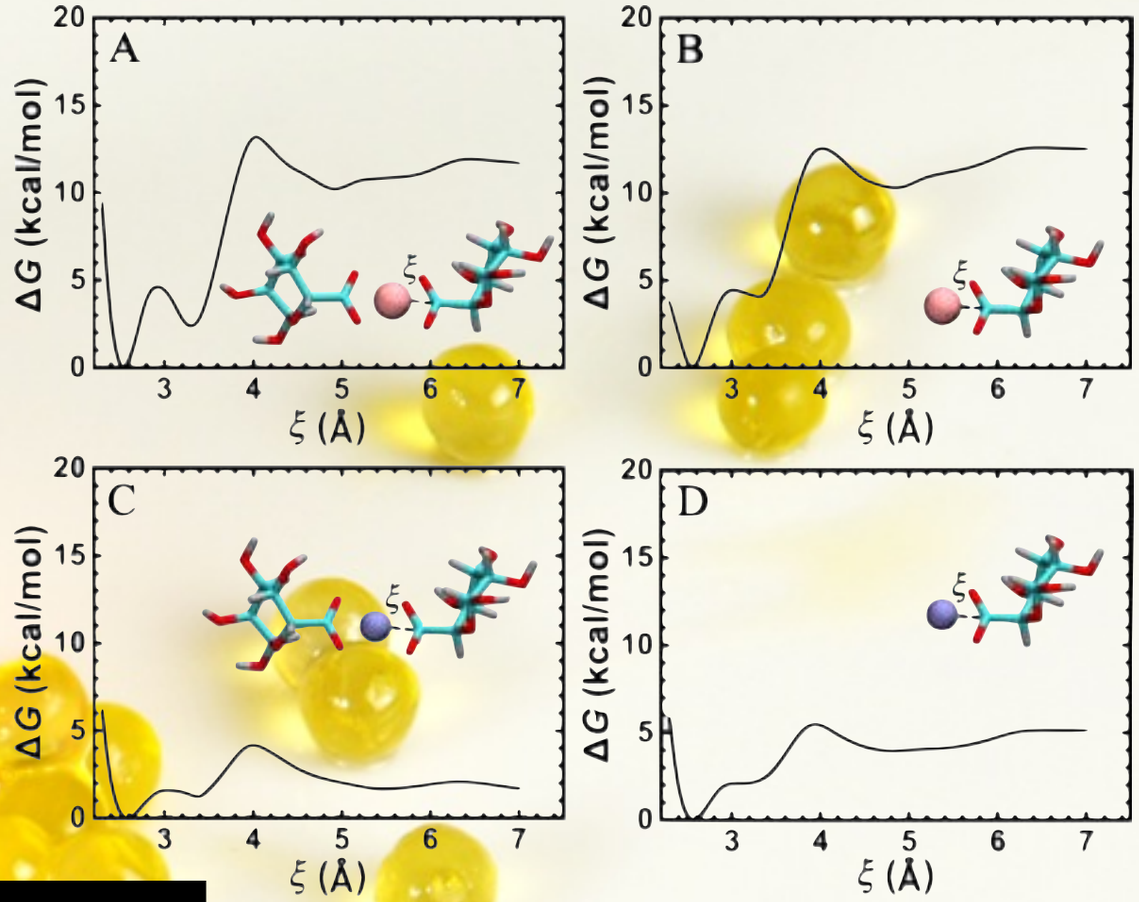
Control simulation: Folding in bulk water.

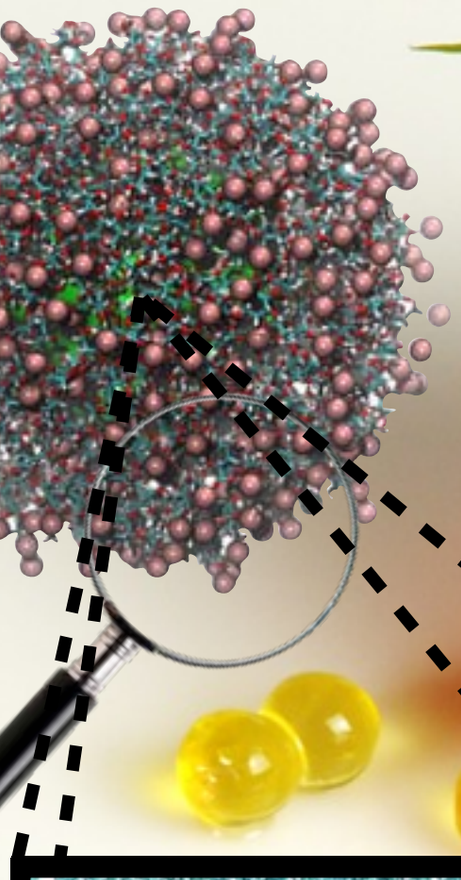
Mingarro, I.; Nilsson, I.; Whitley, P.; von Heijne, G. *BMC Cell Biol.* **2000**, *1*, 3

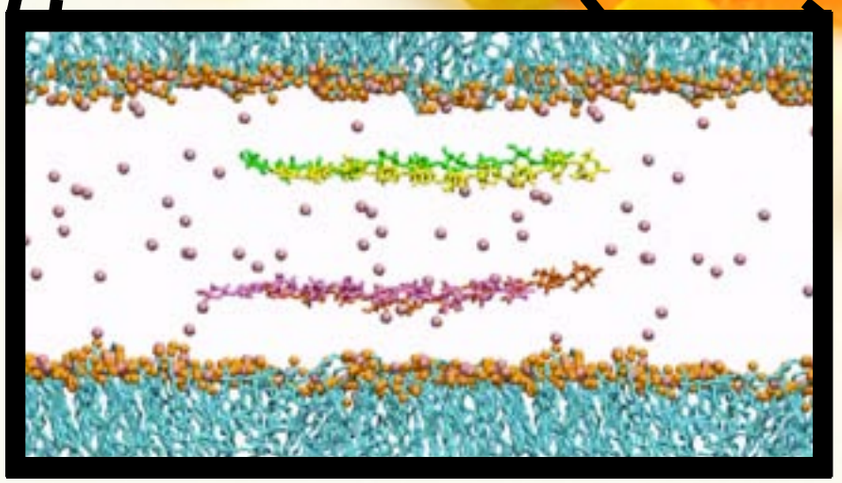
Lu, J.; Deutsch, C. *Nat. Struct. Mol. Biol.* **2005**, *12*, 1123-1129.

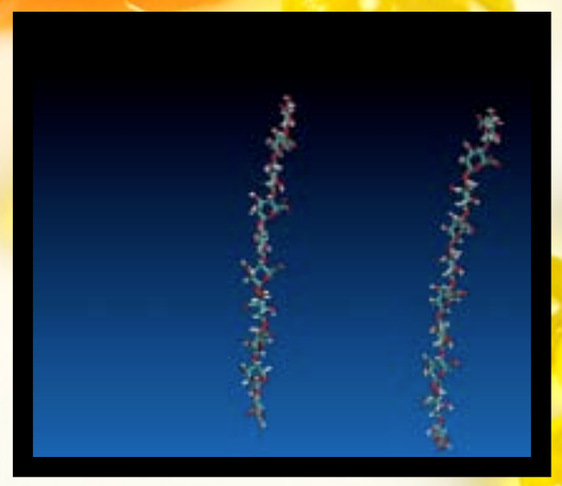
Gumbart, J. C.; Chipot, C.; Schulten, K. *J. Am. Chem. Soc.* **2011**, *133*, 7602-7607

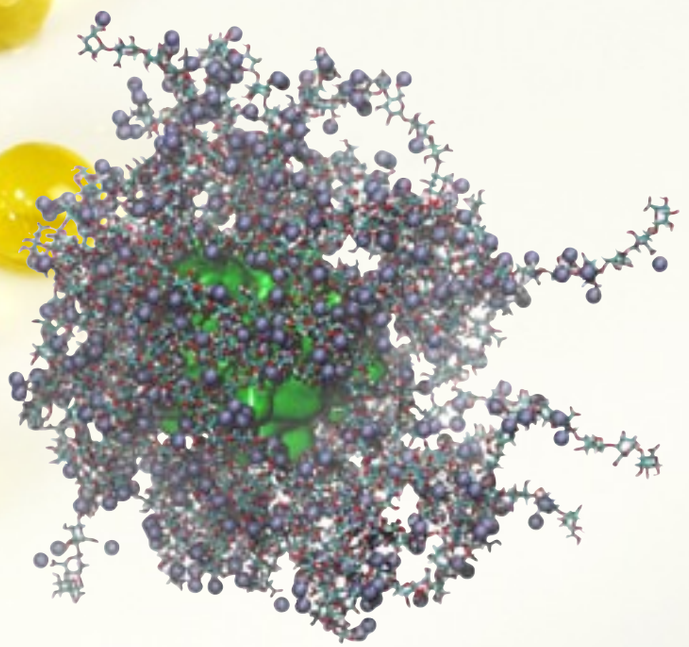
FREE-ENERGY CALCULATIONS AND AVANT-GARDE CUISINE











Fu, H.; Liu, Y.; Adrià, F.; Shao, X.; Cai, W.; Chipot, C. *J. Phys. Chem. B* **2014**, *118*, 11747-11756

Halford, B. *CE&N* **2014**, *92*, 35-36

INTRODUCTION

The unbridled race for longer and larger simulations
What is the best method for a given problem?

ALCHEMICAL FREE-ENERGY CALCULATIONS

- A tool to address host-guest chemistry questions
- Good practices, guidelines and recommendations
- The long-standing protein-ligand problem

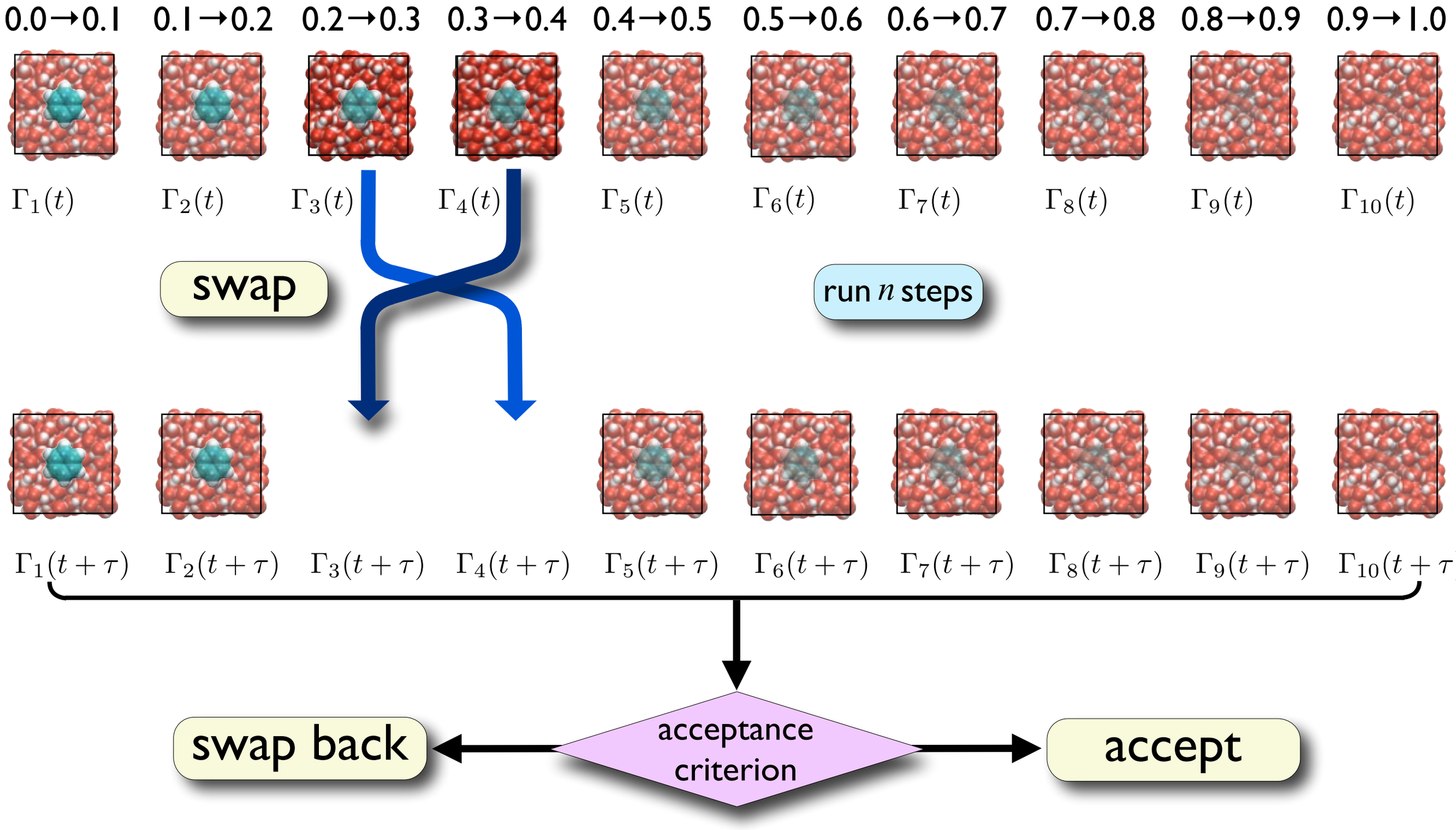
GEOMETRICAL FREE-ENERGY CALCULATIONS

- What is a good reaction-coordinate model?
- A host of methods to measure free-energy changes
- Potentials of mean force and transport phenomena
- Potentials of mean force and recognition and association phenomena
- What about non-equilibrium work computer experiments?

ONGOING CHALLENGES AT THE FRONTIERS OF FREE-ENERGY CALCULATIONS

CONCLUDING REMARKS AND QUESTIONS

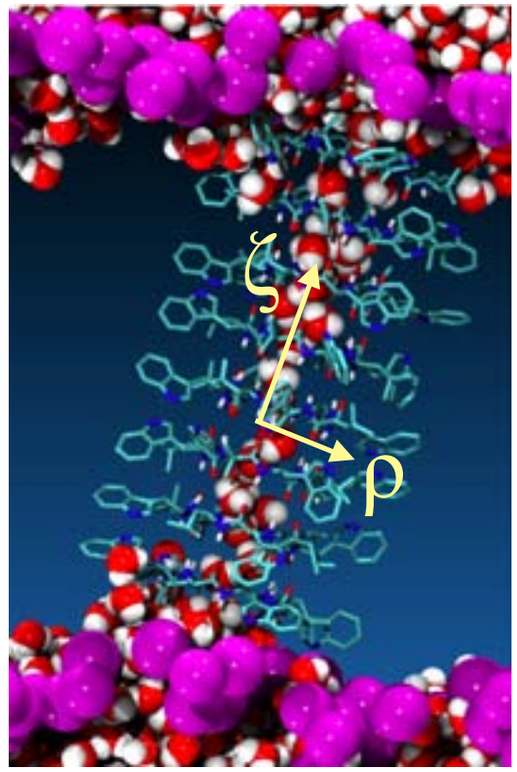
TOWARDS ERGODIC SAMPLING



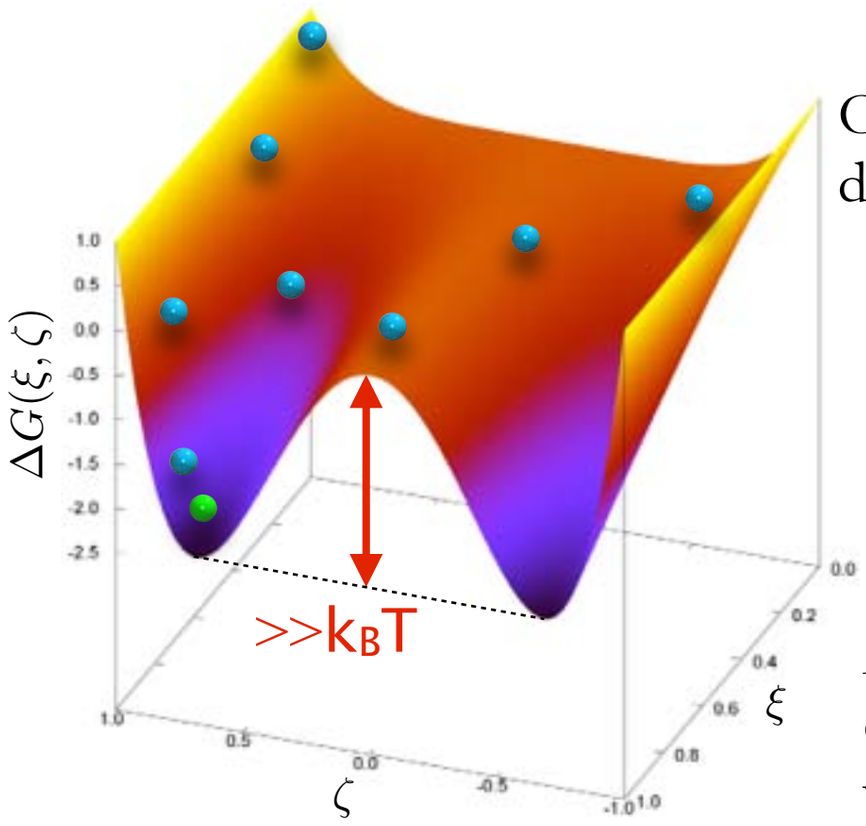
Woods, C. J.; Essex, J. W.; King, M. A. *J. Phys. Chem. B* **2003**, *107*, 13703-13710
Jiang, W.; Hodoscek, M.; Roux, B. *J. Chem. Theory Comput.* **2009**, *5*, 2583-2588
Jiang, W.; Phillips, J.; Huang, L.; Fajer, M.; Meng, Y.; Gumbart, J. C.; Luo, Y.; Schulten, K.; Roux, B. *Comput. Phys. Comm.* **2014**, 185, 908-916

TOWARDS ERGODIC SAMPLING

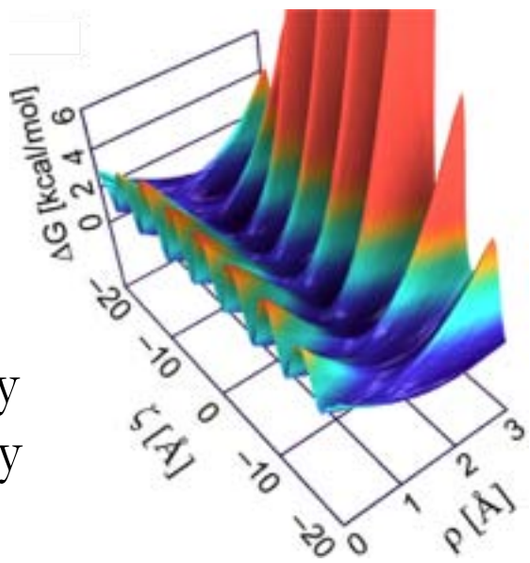
Non-ergodicity scenarios are often related to hidden barriers in orthogonal space.



Such non-ergodicity scenarios generally arise from too naïve a view of what the reaction coordinate is and how we represent it in a low-dimensional collective-variable space.



One possible remedy consists in increasing the dimensionality of the transition coordinate.



Another remedy consists in exploiting the ability of large computer architectures to handle many walkers concomitantly.

Minoukadeh, K.; Chipot, C.; Lelièvre, T. *J. Chem. Theor. Comput.* **2010**, *6*, 1008-101
Comer, J.; Roux, B.; Chipot, C. *Mol. Sim.* **2014**, *40*, 218-228
Comer, K.; Phillips, J.; Schulten, K.; Chipot, C. *J. Chem. Theor. Comput.* **2014**
Zheng, L.; Chen, M.; Yang, W. *Proc. Natl. Acad. Sci. USA.* **2008**, *105*, 20227-20232

INTRODUCTION

The race for longer and larger simulations
What is the best method for a given problem?

ALCHEMICAL FREE-ENERGY CALCULATIONS

- A tool to address host-guest chemistry questions
- Good practices, guidelines and recommendations
- The long-standing protein-ligand problem

GEOMETRICAL FREE-ENERGY CALCULATIONS

- What is a good reaction-coordinate model?
- A host of methods to measure free-energy changes
- Potentials of mean force and transport phenomena
- Potentials of mean force and recognition and association phenomena
- What about non-equilibrium work computer experiments?

ONGOING CHALLENGES AT THE FRONTIERS OF FREE-ENERGY CALCULATIONS

CONCLUDING REMARKS AND QUESTIONS

WHAT DO I NEED FOR THIS WORKSHOP ?

- An up-to-date version of NAMD

- An up-to-date version of VMD, featuring the ParseFEP plugin

- The bundle of hands-on exercises, which can be downloaded from:

```
http://www.ks.uiuc.edu/Training/Tutorials
```