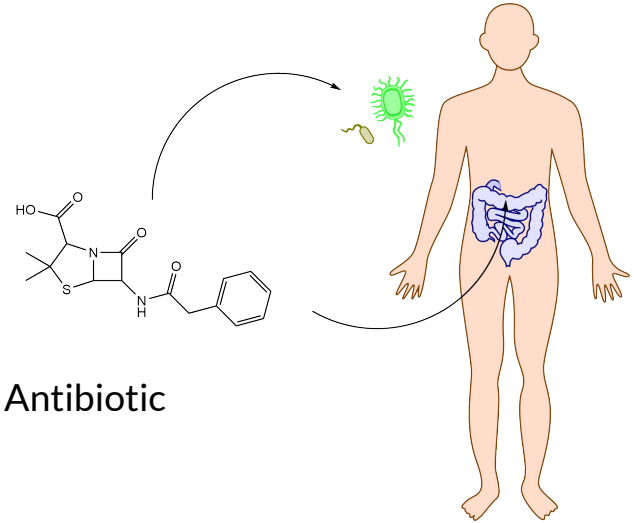
Anchoring Synthetic DNA to a Membrane:
A Novel Application of Free Energy Perturbations

Ezry Santiago
Seminar Presentation
November 9, 2021

Motivation: Drug Delivery

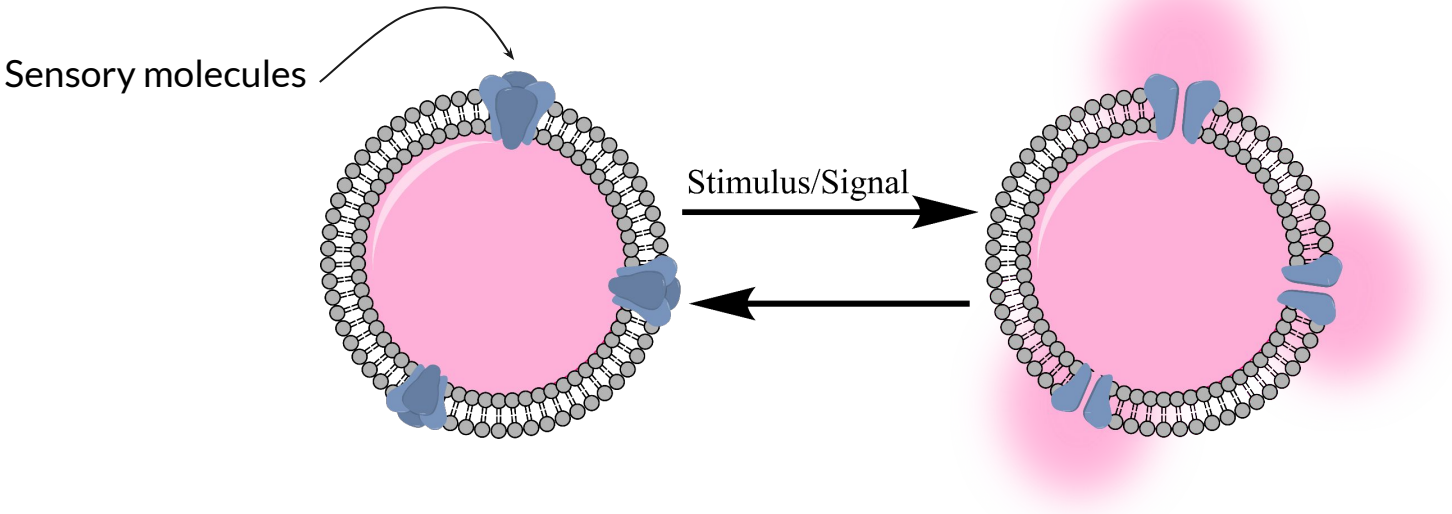
The Problem of Drug Delivery

- The distribution of a drug in the body is hard to control.
- Non-specific delivery leads to side-effects.



Antibiotic

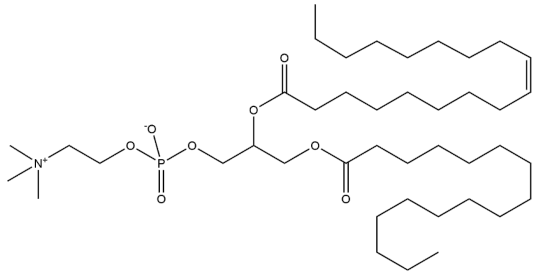
Simple Drug-Delivering Vesicle



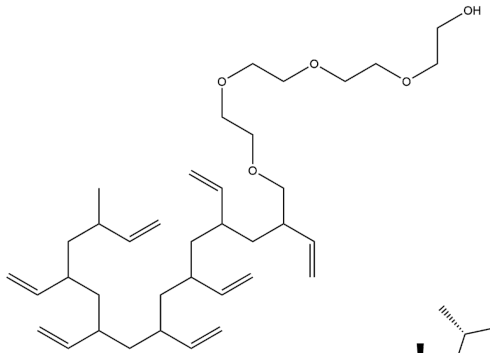
Bottom-Up Design

Polar Lipids (and other amphipaths)

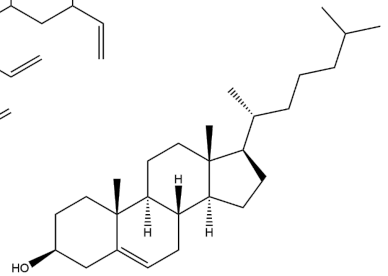
- Obvious choice for membranes
- Cheap
- Easy to manipulate
- Stable and biocompatible



POPC



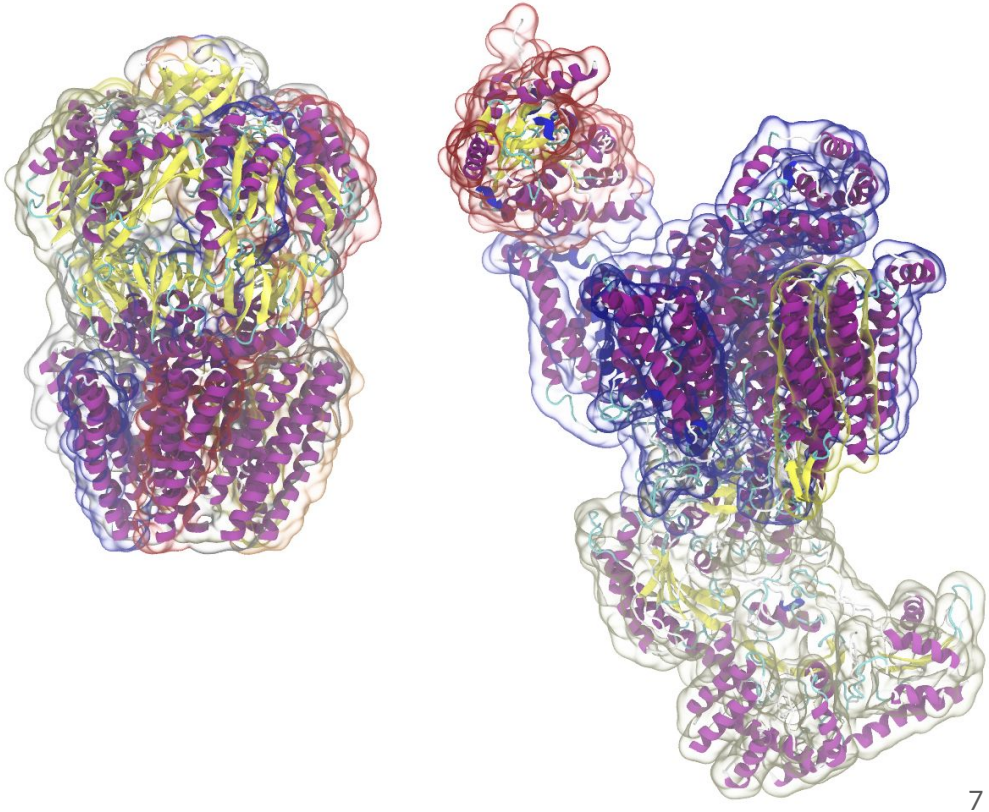
Polymer



Cholesterol

Proteins?

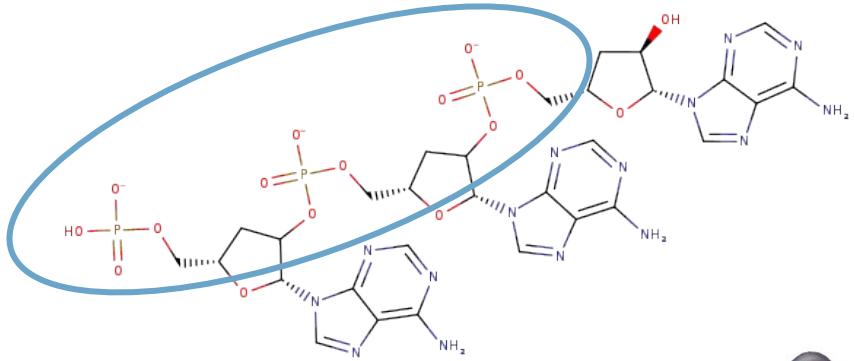
- Pros:
  - Structural and functional
  - Diverse conjugation chemistry
- Cons:
  - Difficult to design
  - Expensive to prototype

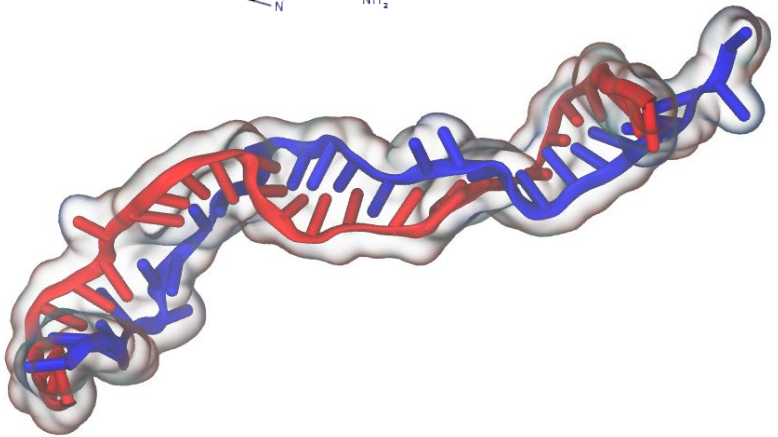


Structures: 2OAU (left) 3JBR (right)

DNA nanotechnology

- Pros:
  - Tractable design due to base-pairing
  - Both structural and functional
  - Diverse conjugation chemistry
- Cons:
  - Less scalable
  - Polyanion



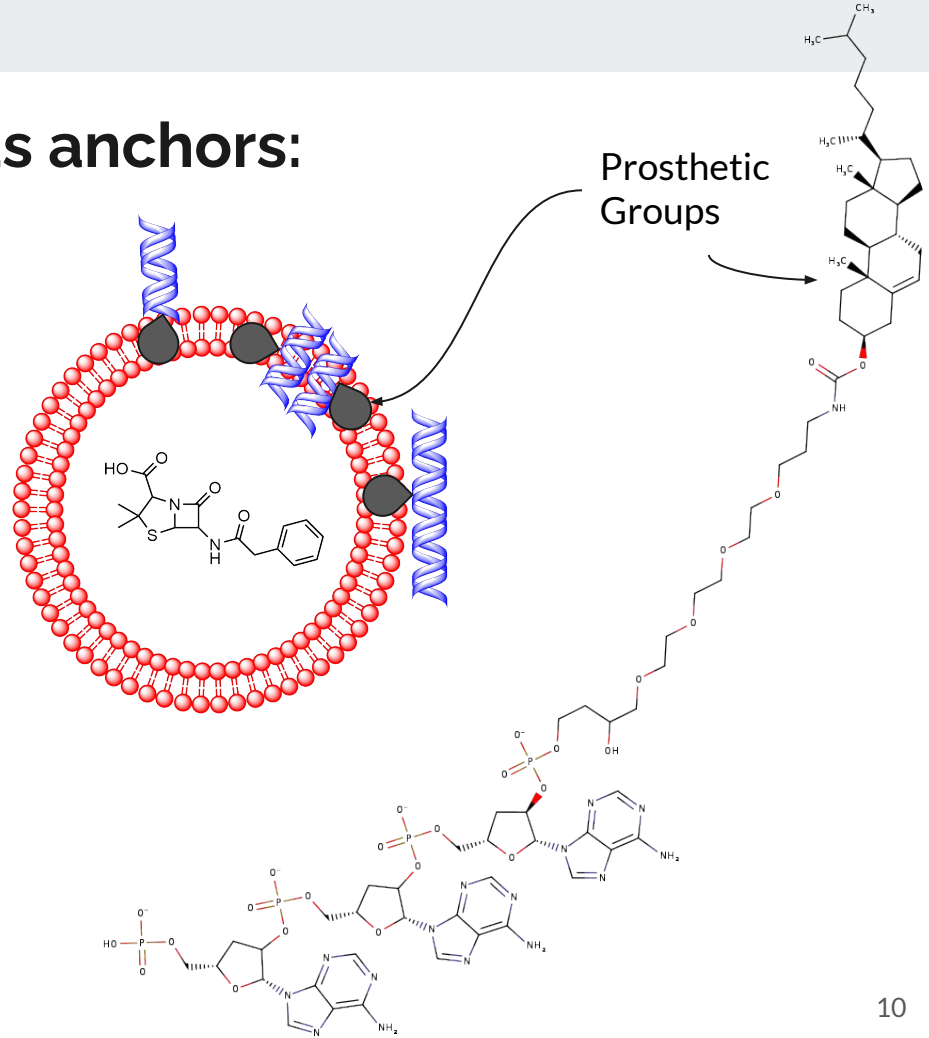


A Problem:

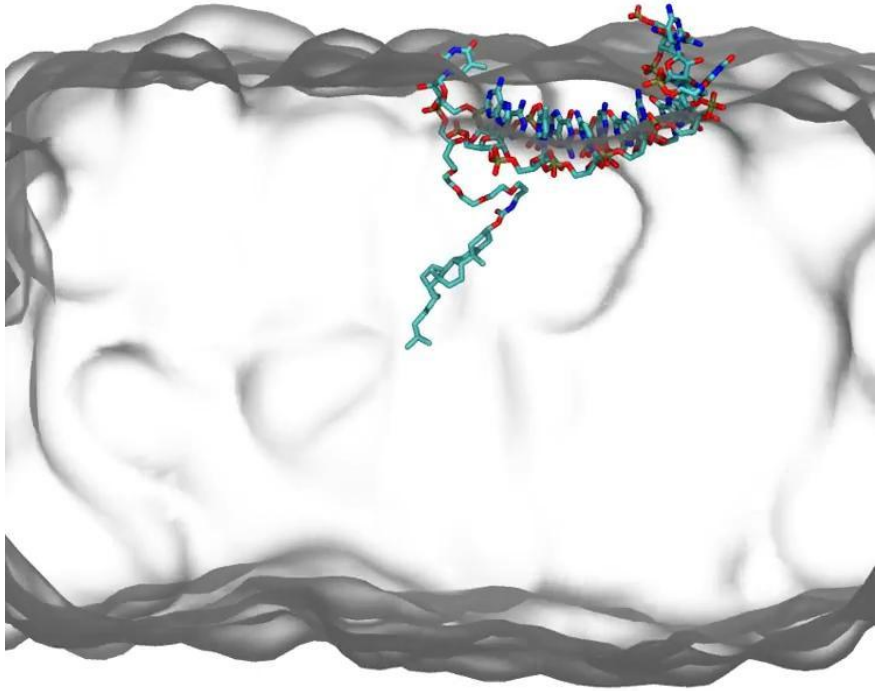
How can we get DNA to interact stably with the hydrophobic region of a membrane?

Prosthetic groups can act as anchors:

- Chemical interfaces between DNA and other molecules



DNA Patched to a Membrane (MD Simulation)



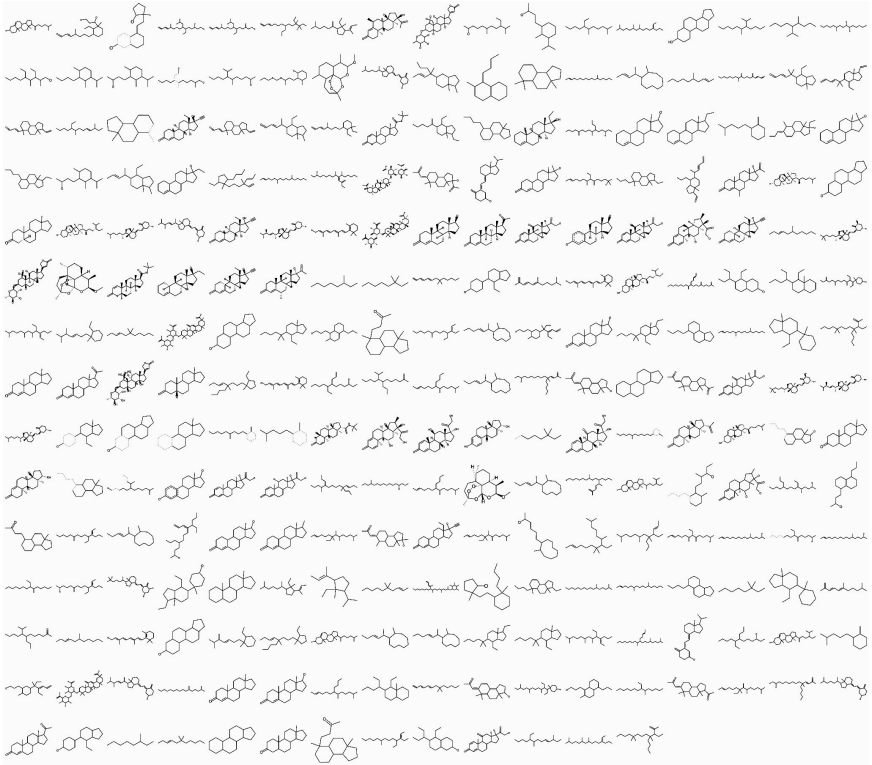
- Nitrogen
- Oxygen
- Carbon
- Phosphorous
- Lipids

Water not shown

The question:

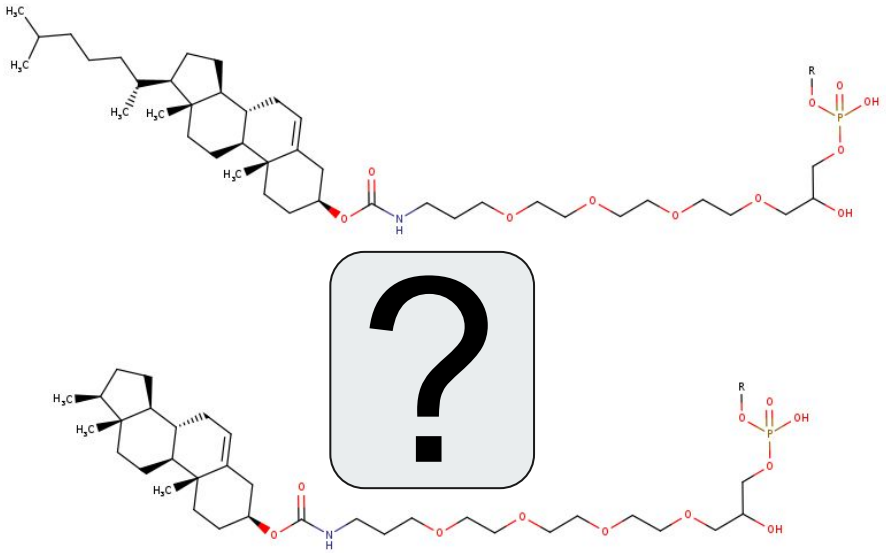
Which prosthetic group stabilizes the DNA-Membrane complex most?

Current Problem: Which prosthetic group is optimal?



a small selection of possible prosthetic groups
(sterols and sphingosine derivatives)

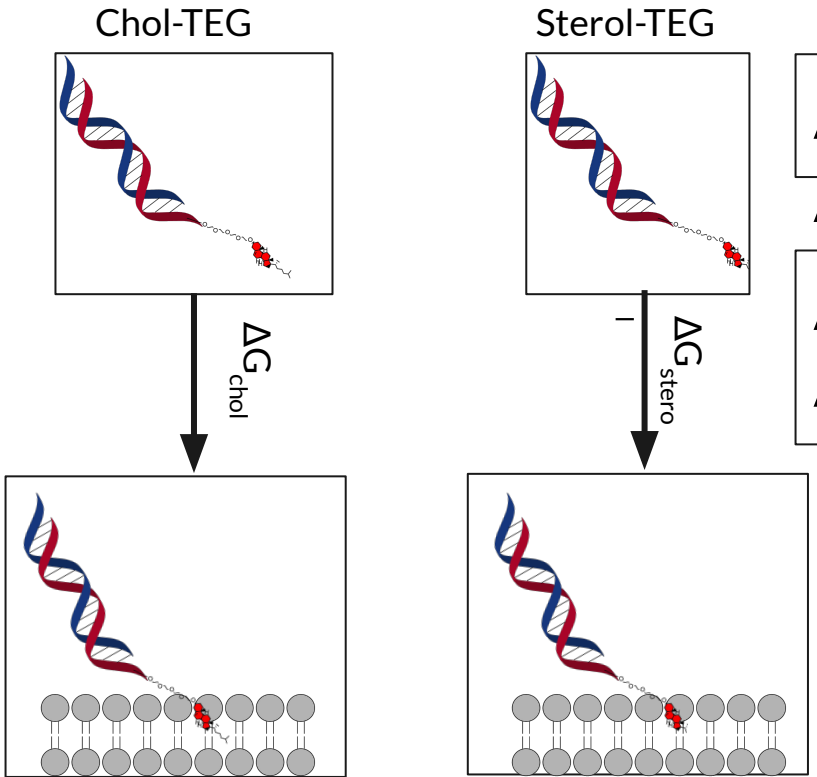
Which is Better: CholTEG or SterolTEG?



Our Approach:

Comparing Free Energies of Insertion

Free Energy Comparisons



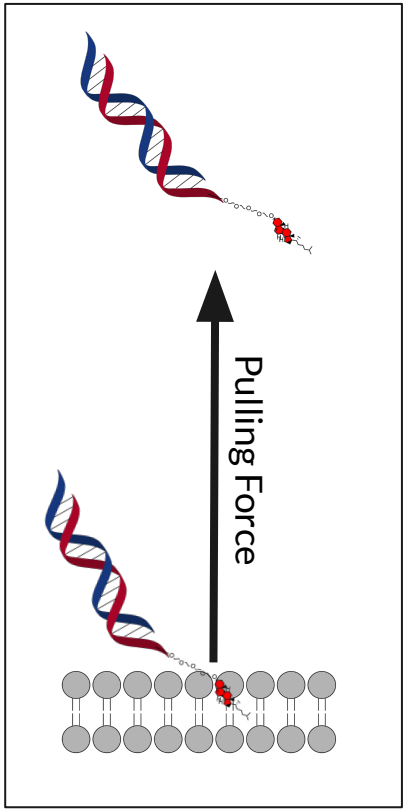
$$\Delta\Delta G_{chol \rightarrow sterol} = \Delta G_{sterol} - \Delta G_{chol}$$

ΔΔG<0 Chol is more stable
ΔΔG>0 Sterol is more stable

Free Energy Calculations Option 1: Spatial Method(s)

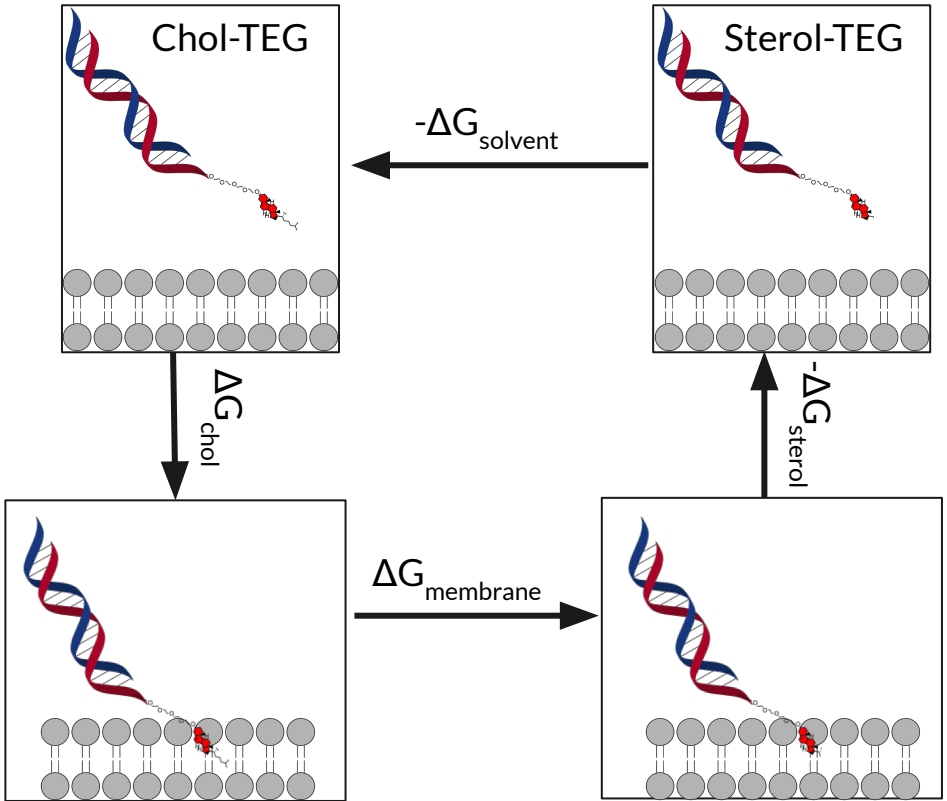
Example:

1. Push/pull the prosthetic group in/out of the membrane
2. Measure the resulting reactive forces to estimate the change in F over time
3. Sum the changes over dt to get the total
4. Repeat for each prosthetic group



Images made with ChemDraw

Free Energy Comparisons: Thermodynamic Cycle



$$\Delta G_{chol} + \Delta G_{membrane} - \Delta G_{sterol} - \Delta G_{solvent} = 0$$

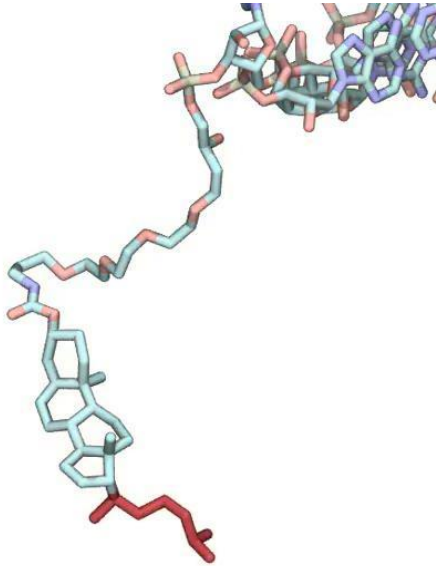
$$\Delta\Delta G_{chol \rightarrow sterol} = \Delta G_{sterol} - \Delta G_{chol}$$

$$\Delta\Delta G_{chol \rightarrow sterol} = \Delta G_{membrane} - \Delta G_{solvent}$$

Getting $\Delta G_{membrane}$ and $\Delta G_{solvent}$: Alchemical Free Energy Calculations (FEP)

$$\underset{a\to b}{\Delta G} = \sum_{\lambda_i=0}^{1} \ln \left\langle \exp\left[-\beta\left[E_{\lambda_{i+1}} - E_{\lambda_i}\right]\right]\right\rangle_{\lambda_i}$$

1. Decompose the two states into N intermediates: $\{\lambda_0,\lambda_1,\lambda_2...\lambda_N\}$
2. Simulate in state λ_i
3. Periodically calculate the internal energy (E) as if we were in state λ_{i+1}
4. Use the Boltzmann distribution to estimate $\Delta G(\lambda_i, \lambda_{i+1})$
5. Sum over all substates



Uses and Challenges of Alchemical Free Energy Perturbations

Uses:

- Traditionally used in drug discovery
- Application to membrane dynamics is very new (<5 years)
- When it works, it’s much more efficient than spatial methods

Challenges:

- Windows can’t be too wide
- Each window must be simulated for a sufficient time

Question:
Will FEP work in this system?

Two Main Criteria:

1. Convergence:

   Does our estimate stabilize over simulation time?

2. Error:

   How large is the remaining error?

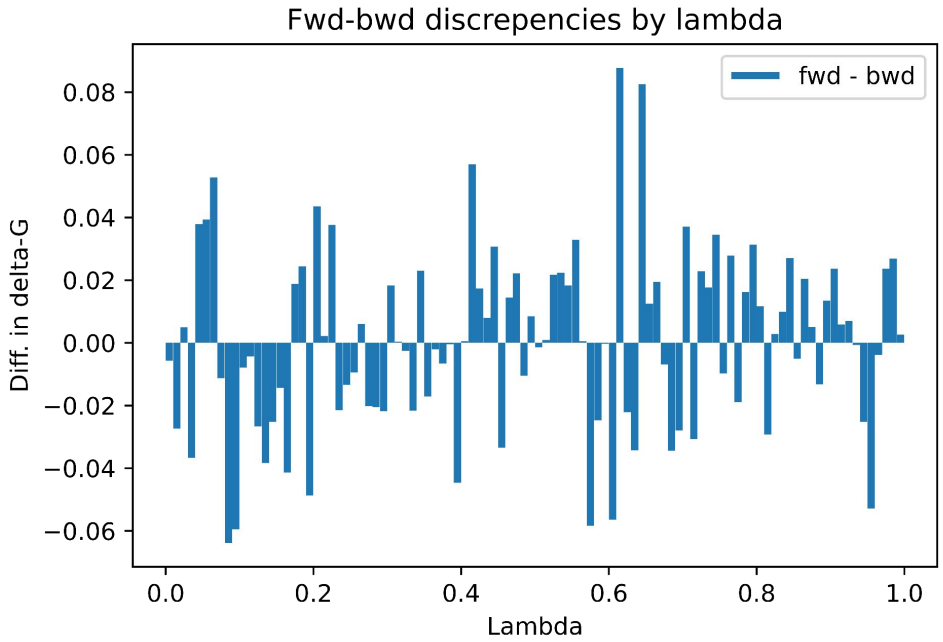
Hysteresis Plots and Convergence

$$\underset{a\to b}{\Delta G} = \sum_{\lambda_i=0}^{1} \ln \left\langle \exp\left[-\beta\left[E_{\lambda_{i+1}} - E_{\lambda_i}\right]\right]\right\rangle_{\lambda}$$

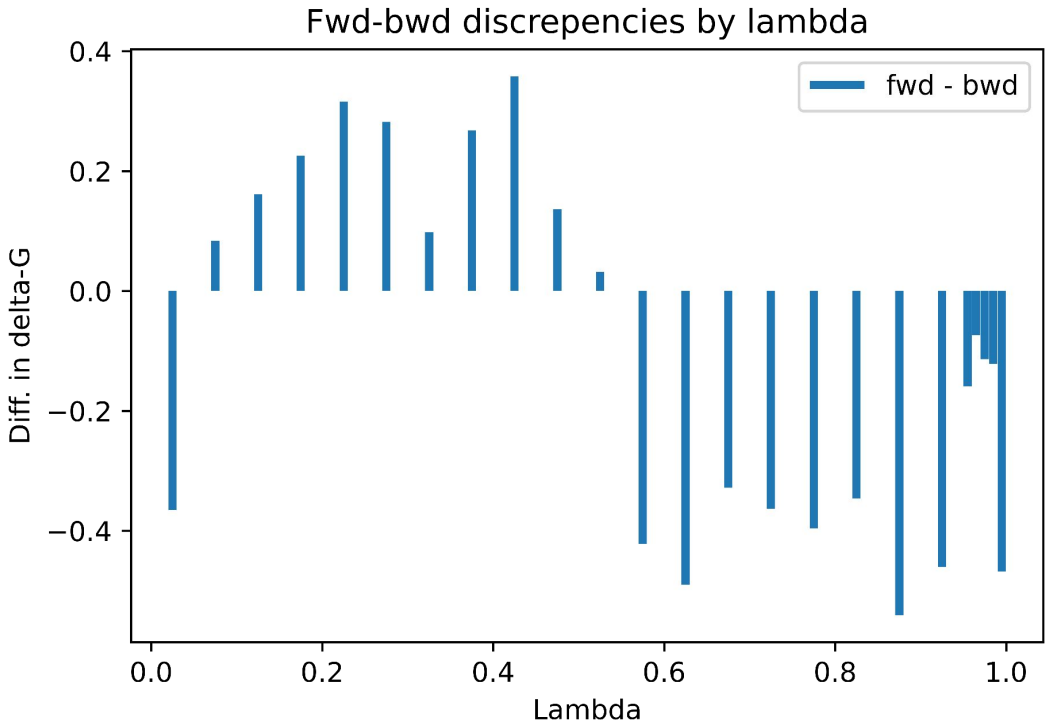
For every pair λ_i and λ_{i+1}

With sufficient simulation time:
$\Delta G(\lambda_i, \lambda_{i+1}) = \Delta G(\lambda_{i+1}, \lambda_i)$

I.e. The calculation **converges**



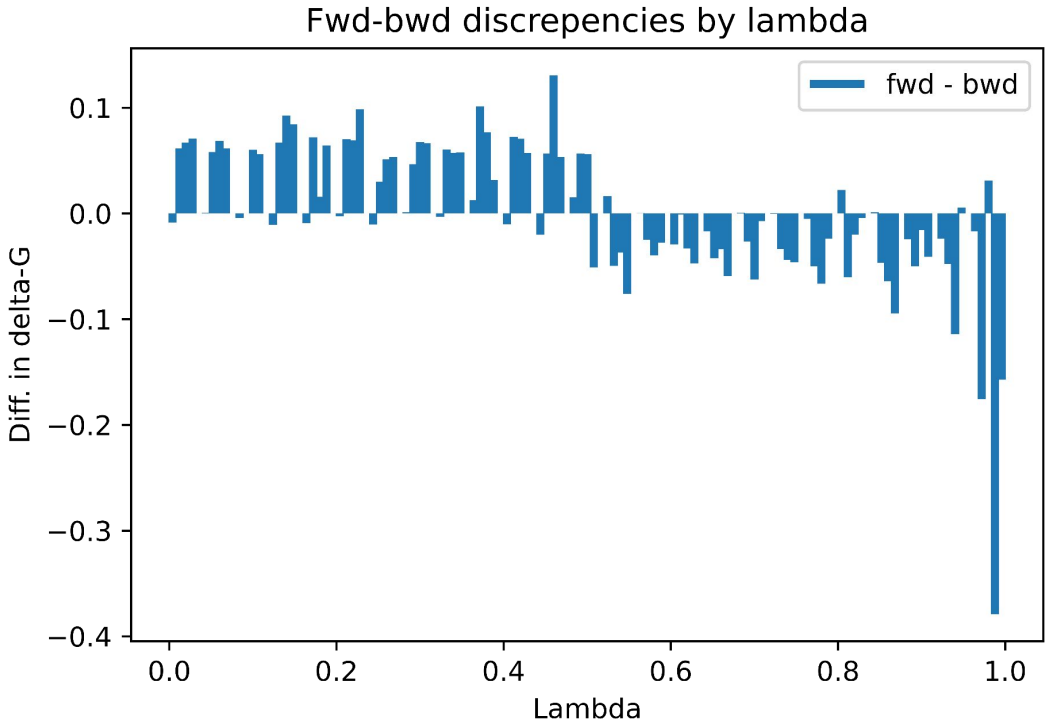
First Attempt (July)



Issues:

1. Differences are large
2. Sign switch near 0.5

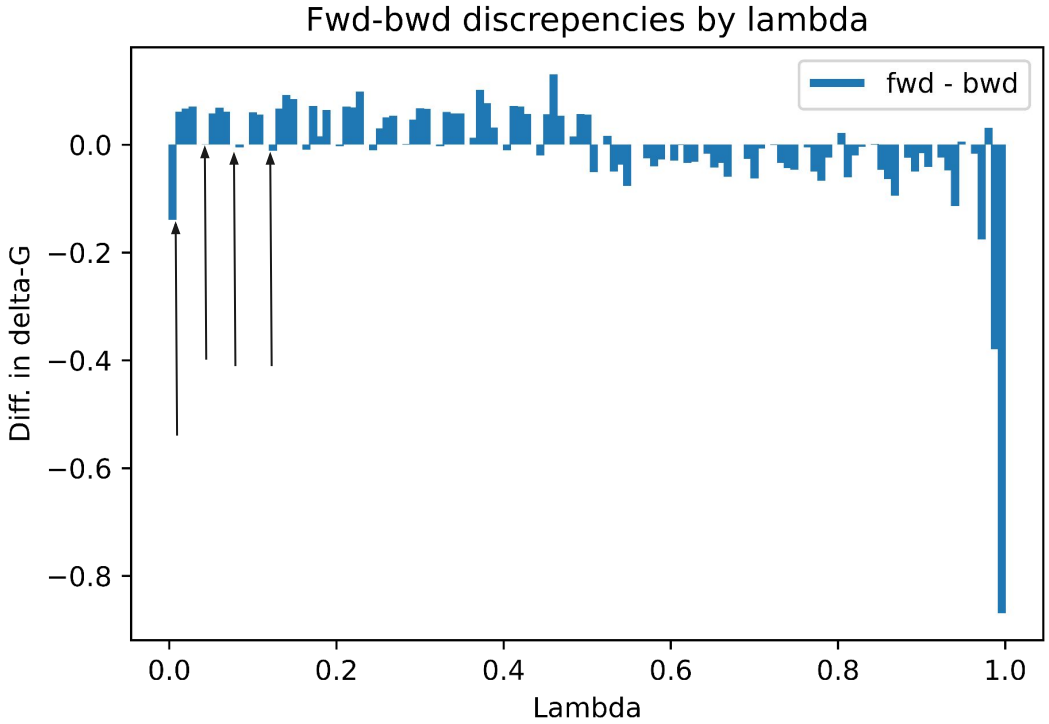
Second Attempt



Issues:

1. Differences are large
2. Sign switch near 0.5
3. Stubborn difference at 1

Penultimate Attempt



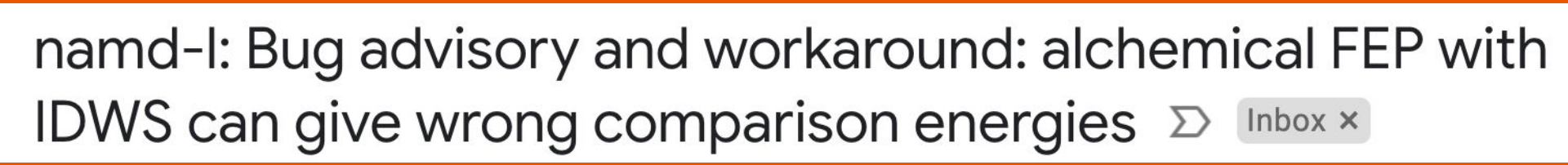
We tried everything

Issues:

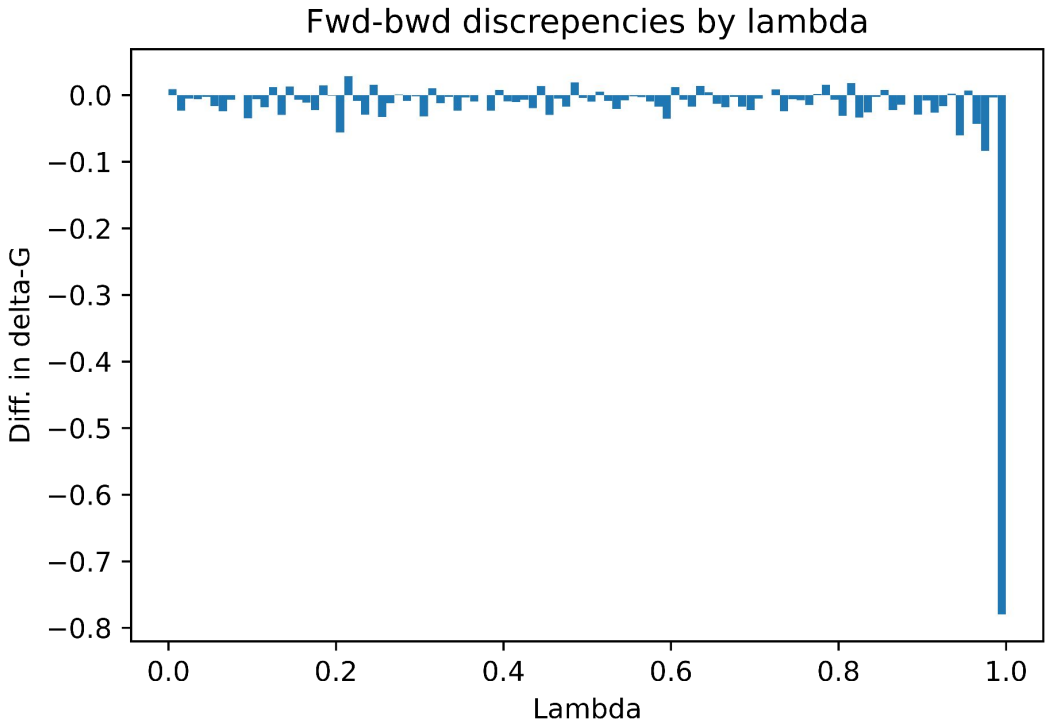
1. Differences are large
2. Sign switch near 0.5
3. Stubborn difference at 1

There was a bug in NAMD

With implications for many other labs



Latest Results (still running, preliminary)

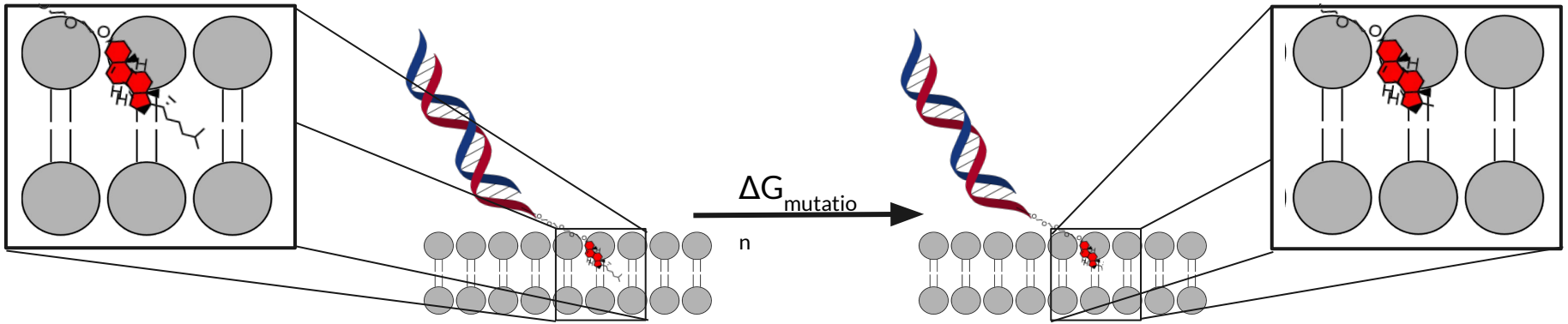


Issues:

1. ~~Differences are smaller still~~
2. ~~Sign switch near 0.5~~
3. Stubborn difference at 1

Sterol is more stable in the membrane

$$\Delta G_{mutation} = -1.6 \pm 0.6 \text{ kcal/mol}$$



Summary

Applicability	Cholesterol vs Sterol	Improvements	More Prosthetic Groups
We have demonstrated that free energy perturbation calculations converge (work) for these systems	We now know that a sterol-TEG is marginally more stable than chol-TEG in the membrane Need to calculate $\Delta G_{solution}$	Calculations still take several days. This can be improved.	Selection of more prosthetic groups is already underway

Acknowledgements

Advisors:

- Dr. Grace Brannigan — Primary Advisor
- Dr. Jinglin Fu — Secondary Advisor
- Dr. Julie Griepenburg — Collaborator
- Dr. Jerome Henin — Collaborator
- Dr. Tom Joseph — Collaborator

Lab Members:

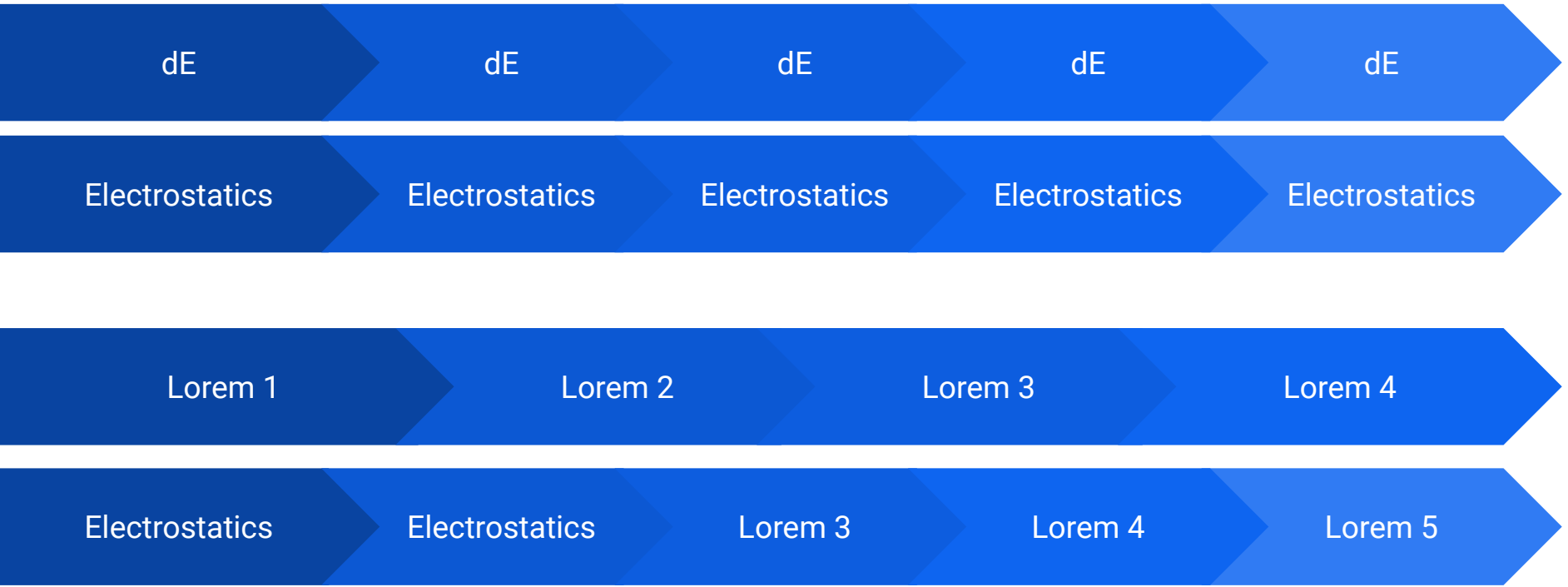
- Jesse Sandberg, MS
- Connor Pitman
- Jahmal Ennis

Resources:

- Compute Resources: Office of Advanced Research Computing (OARC)
- Funding: Department of Defence (DoD)

How do biological systems deliver metabolites across the body?

The biological approach (e.g. Oxygen Transport):


dE
dE
dE
dE
dE
Electrostatics
Electrostatics
Electrostatics
Electrostatics
Electrostatics
Lorem 1
Lorem 2
Lorem 3
Lorem 4
Electrostatics
Electrostatics
Lorem 3
Lorem 4
Lorem 5

Outline

Introduction:

- Big-picture - moving and manipulating small groups of molecules
- Biological inspiration - vesicles and associated proteins
- Our toolbox - DNA, chemical modifications, and synthetic vesicles

Problem statement: optimizing prosthetic groups

Approach: Collect and organize candidates

- Compare the free energies of solvation for various prosthetic groups
  - Energy components

Methods: Free energy perturbation (compare and contrast with LogP)

Results: Convergence and hysteresis?

Applicability of FEP:

		Proposed	Tested	Published
0	Equations of state	✔	✔	1954
1	Calculation of partition coefficients (ΔG_solvation)	✔	✔	✔
2	Ligand Binding in Dilute Aqueous Solution (ΔG_binding)	✔	✔	✔
3	Ligand Binding in Heterogeneous, Non-Dilute Systems	✔	✔	2018
4	Patching on Membranes	✔	This work	✖
5	Insertion of Transmembrane Macromolecules	✔	Upcoming	✖

RW Zwanzig

Salari et al.

Ensuring Convergence and Reducing Error

Our “Best Practices” so far:

1. Sample frequently
    a. ~every 16-20 fs
2. Use double-wide sampling
    a. Simulate intermediate state N while sampling N-1 and N+1
3. Equilibrate thoroughly
    a. Initially ~100ns
    b. ~10ps for each window
4. Use long enough sampling times (~2ns)
5. Use small windows to ensure good overlap (~100 divisions between A and B)

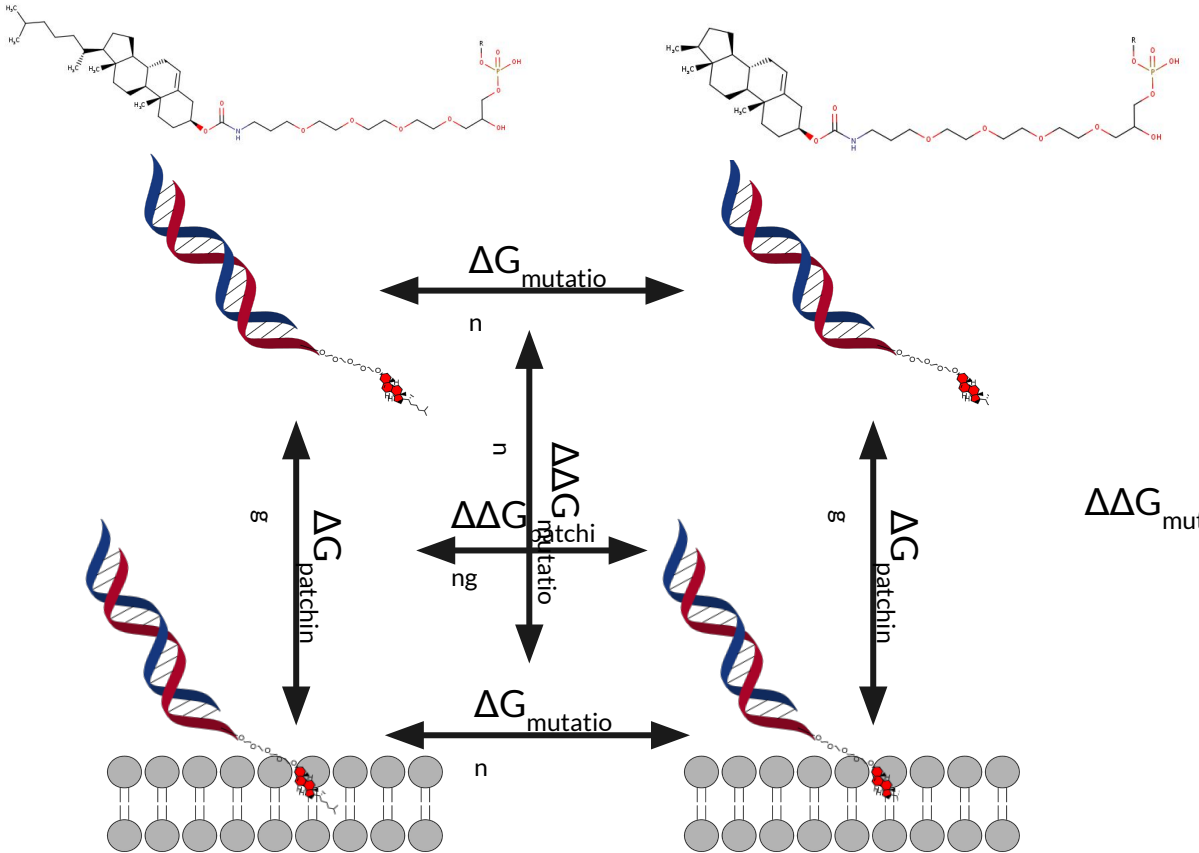
~~Add thermodynamic cycle~~

Add movie of mutation

~~Add slide numbers~~

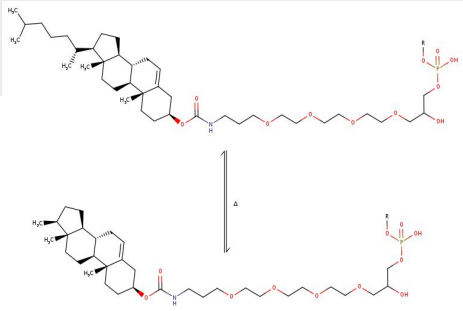
~~Define convergence~~

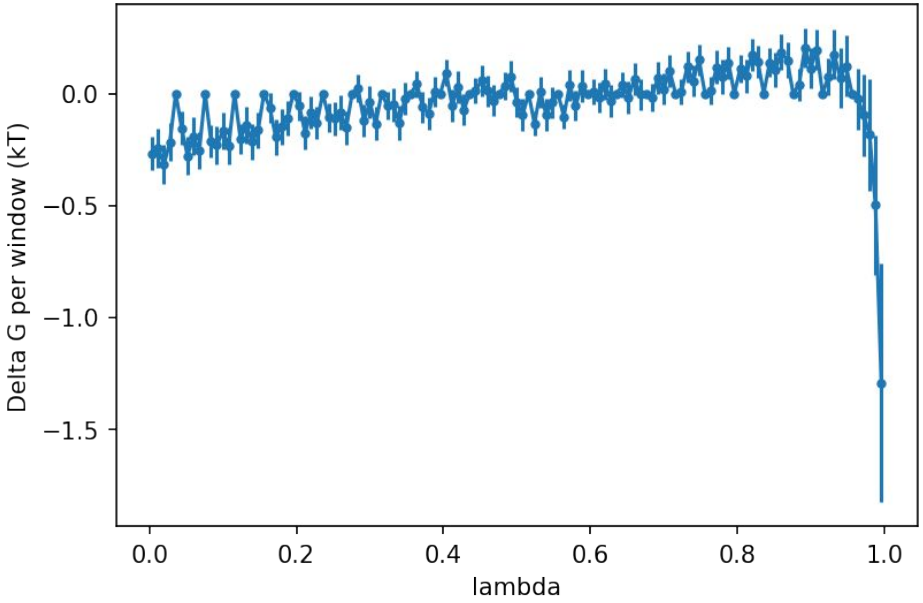
Free Energy Comparisons

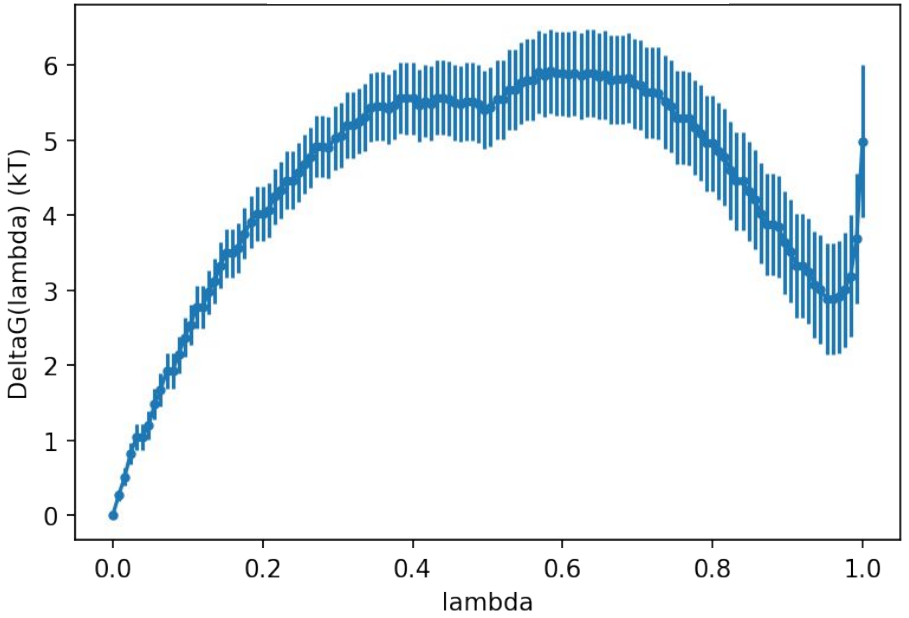


$$\Delta\Delta G_{mutation} = \Delta\Delta G_{patching}$$

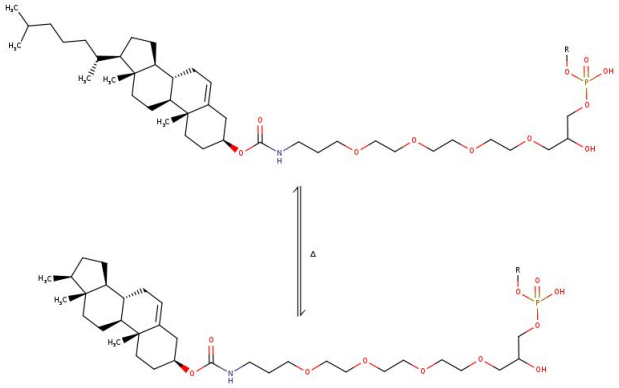
Best estimate so far:







Combining DNA and Membranes



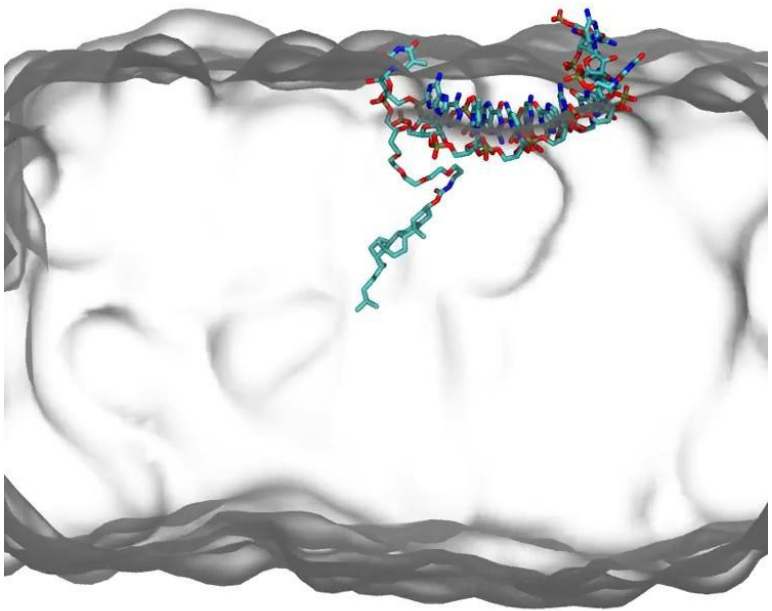
- Nitrogen
- Oxygen
- Carbon
- Phosphorous
- Lipids

Refinement: decompose the tw
intermediates: $\{\lambda_0,\lambda_1,\lambda_2...\lambda_N\}$

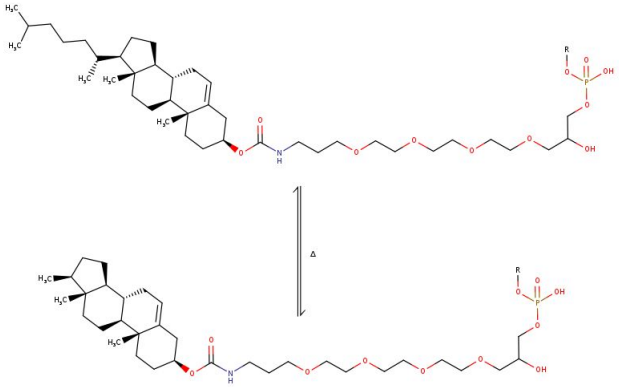
A) Simulate in state A
B) Periodically calculate the internal energy (E) *as if* we were in state B
C) Use the Boltzmann distribution to estimate dF

$$\Delta F(\mathbf{A} \rightarrow \mathbf{B}) = F_{\mathbf{B}} - F_{\mathbf{A}}$$

$$\underset{a \rightarrow b}{\Delta F} = \sum_{\lambda_i=0}^{1} \ln \left\langle \exp\left[-\beta[E_{\lambda_{i+1}} - E_{\lambda_{i+1}}]\right] \right\rangle_{\lambda}$$



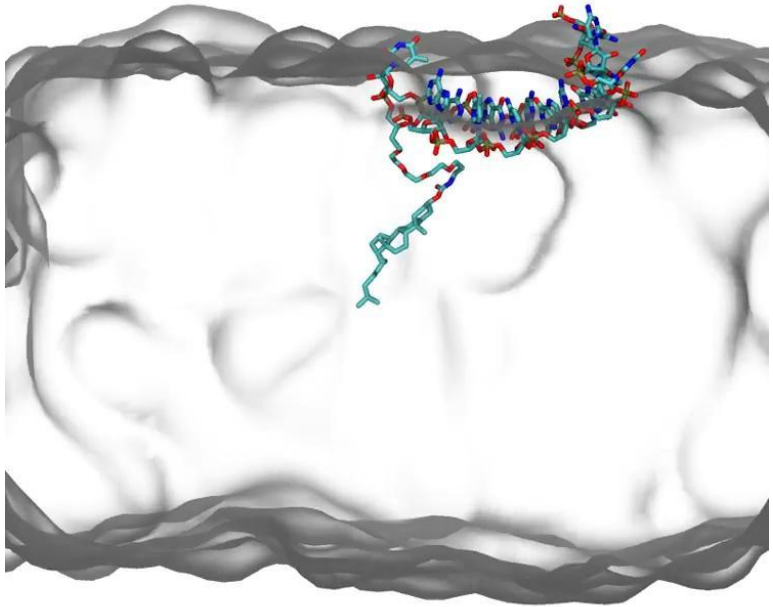
AFEP: a closer look



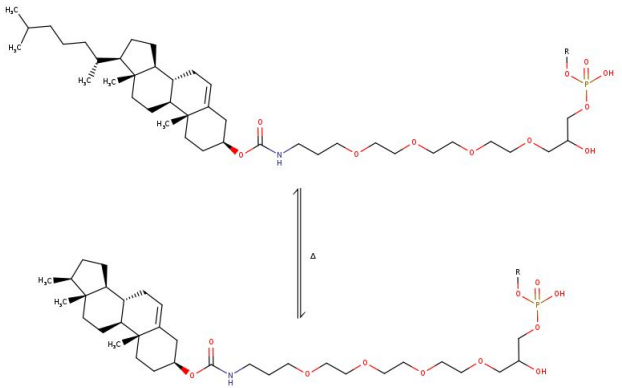
$$\Delta F(\mathbf{A} \rightarrow \mathbf{B}) = F_{\mathbf{B}} - F_{\mathbf{A}}$$

$$\underset{a \rightarrow b}{\Delta F} = \sum_{\lambda_i=0}^{1} \ln \left\langle \exp\left[-\beta[E_{\lambda_{i+1}} - E_{\lambda_{i+1}}]\right] \right\rangle_{\lambda}$$

- Nitrogen
- Oxygen
- Carbon
- Phosphorous
- Lipids



Organization Paradigms



A) by size Maximum Common Structure:

B) by Tanimoto Distance:

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}.$$

Basic cheminformatics:

C) by MCS with tuning by CHARMM parameter distances:

A, but including CHARMM parameters as additional data

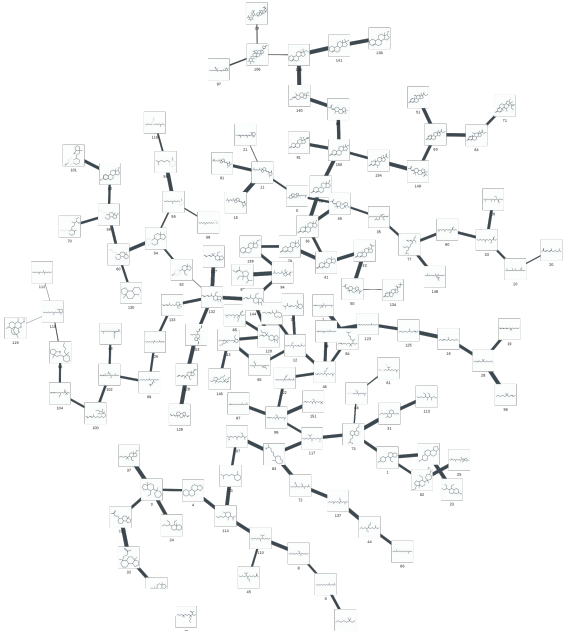
Approach

Organize candidate prosthetic groups

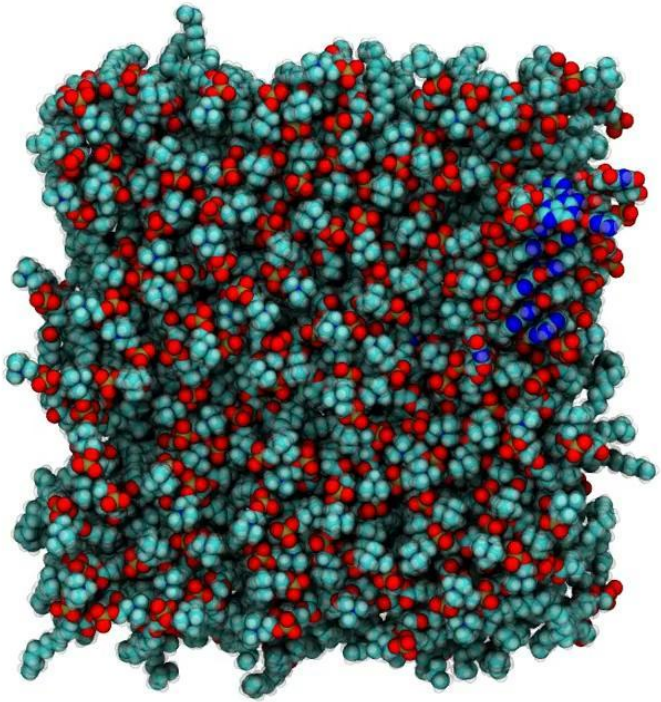
↓

Compute the relative free energies of solvation/insertion for each

↑



Methods: Simulation



- Nitrogen
- Oxygen
- Carbon
- Phosphorous
- Lipids