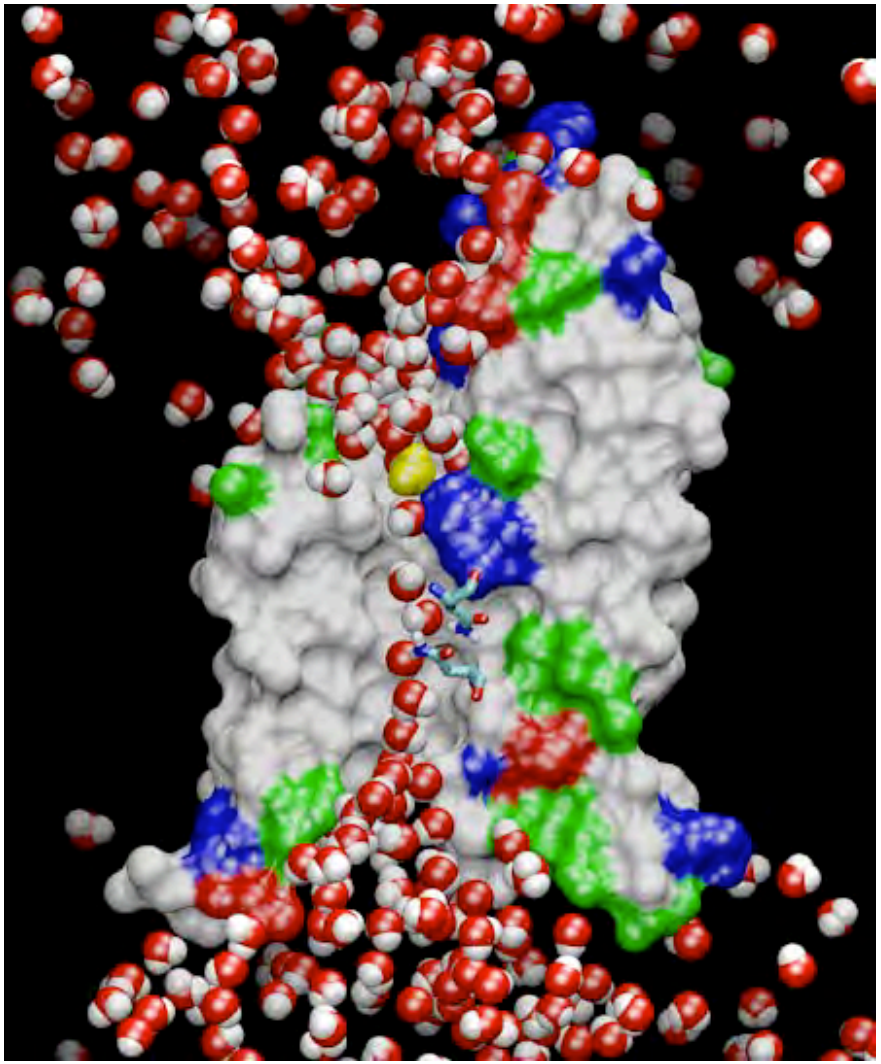
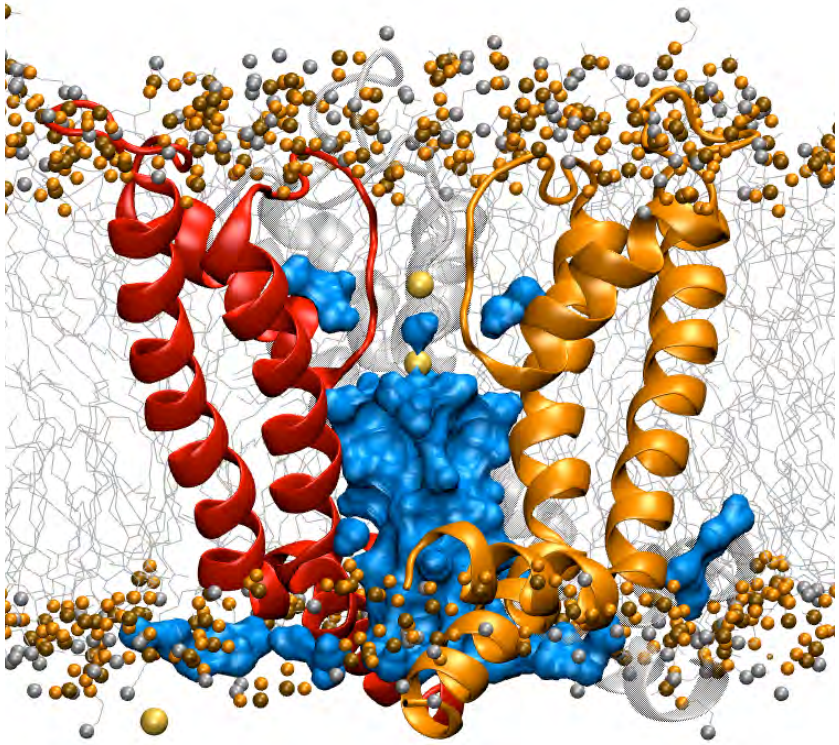
Modeling and Molecular Dynamics of Membrane Proteins



Emad Tajkhorshid

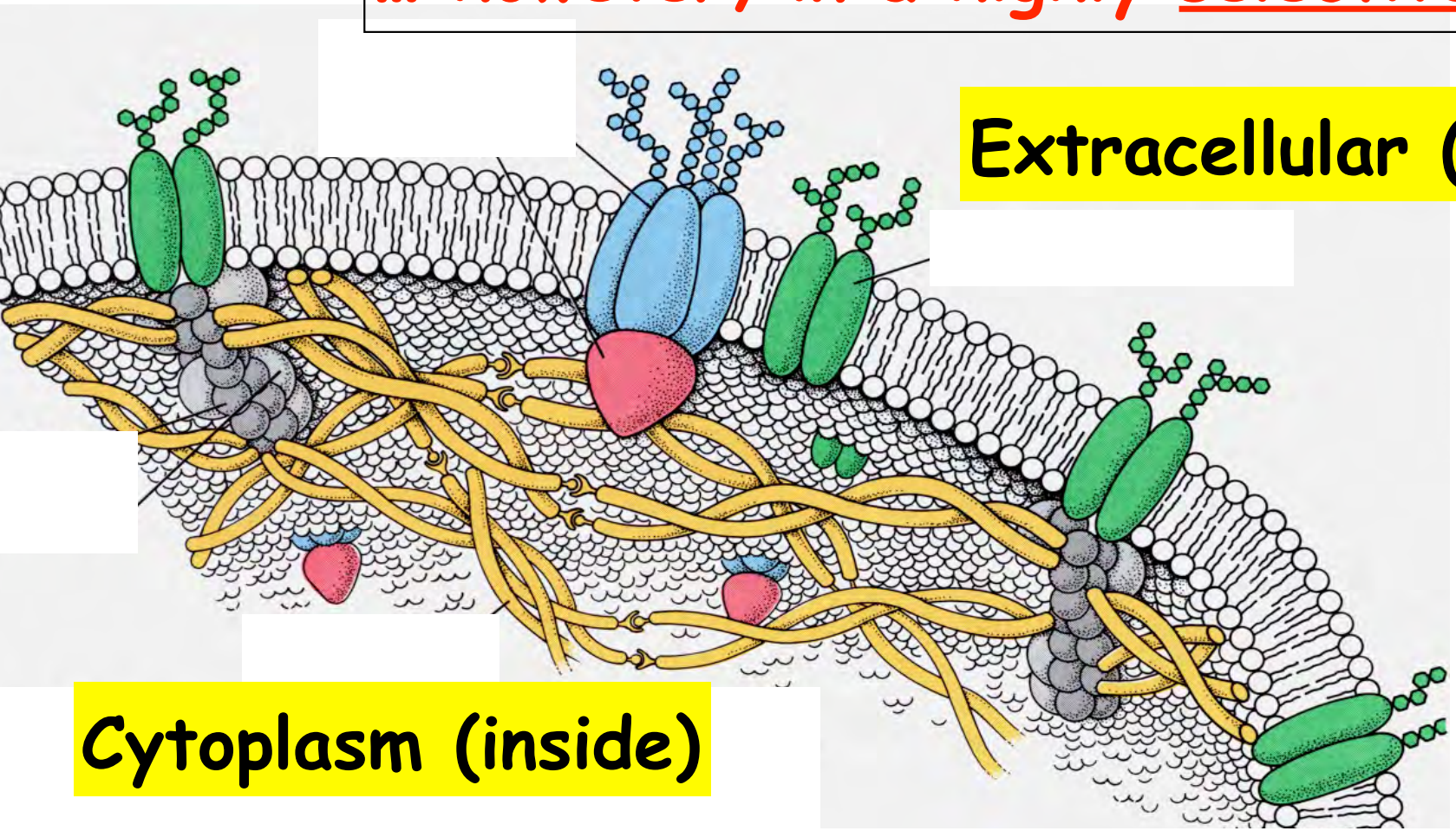
NIH Center for Macromolecular Modeling and Bioinformatics, Beckman Institute, University of Illinois at Urbana-Champaign



Why Do Living Cells Need Membrane

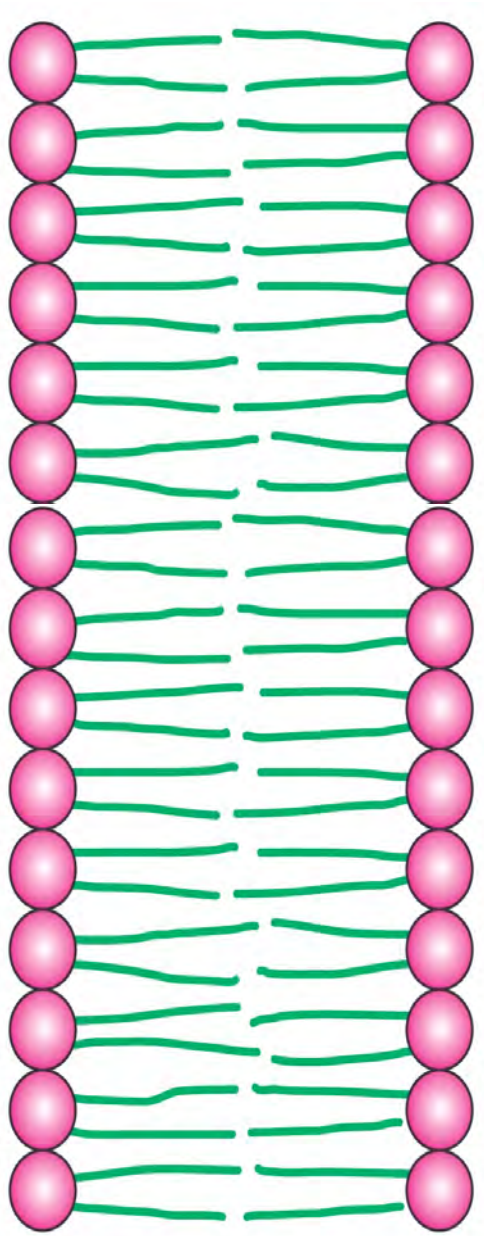
- Living cells also need to exchange materials and information with the outside world

… however, in a highly selective manner.

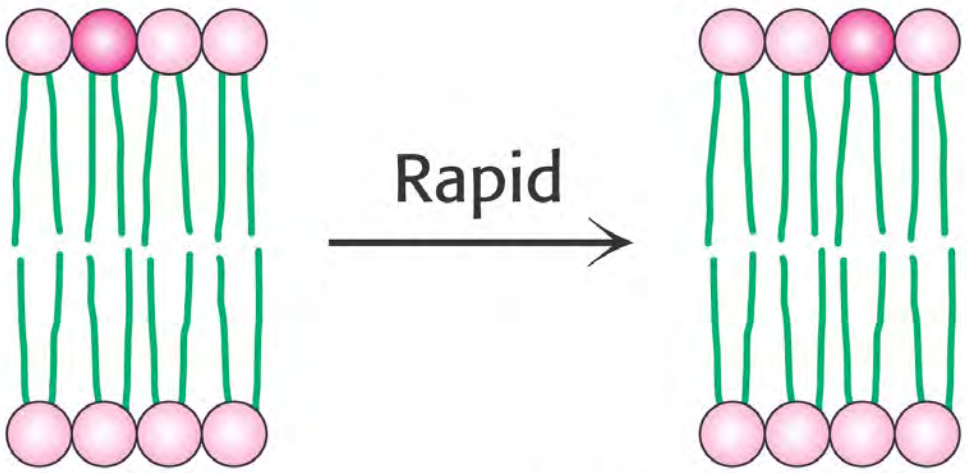


Phospholipid Bilayers Are Excellent Materials For Cell Membranes

- Hydrophobic interaction is the driving force
- Self-assembly in water
- Tendency to close on themselves
- Self-sealing (a hole is unfavorable)
- Extensive: up to millimeters



Lipid Diffusion in a Membrane



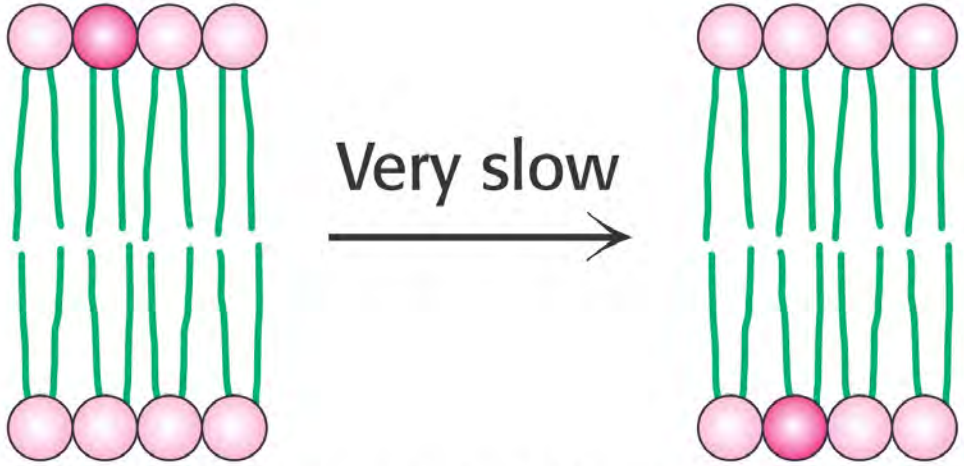
Lateral diffusion

$D_{lip} = 10^{-8}\ cm^2.s^{-1}$

(50 Å in ~ 5 x 10^{-6} s)

$D_{wat} = 2.5 \times 10^{-5}\ cm^2.s^{-1}$

Modeling mixed lipid bilayers!



Tranverse diffusion
(flip-flop)

Once in several hours!

(~ 50 Å in ~ 10^4 s)

~9 orders of magnitude slower

ensuring bilayer asymmetry

Technical difficulties in Simulations of Biological Membranes

- Time scale
- Heterogeneity of biological membranes ☹

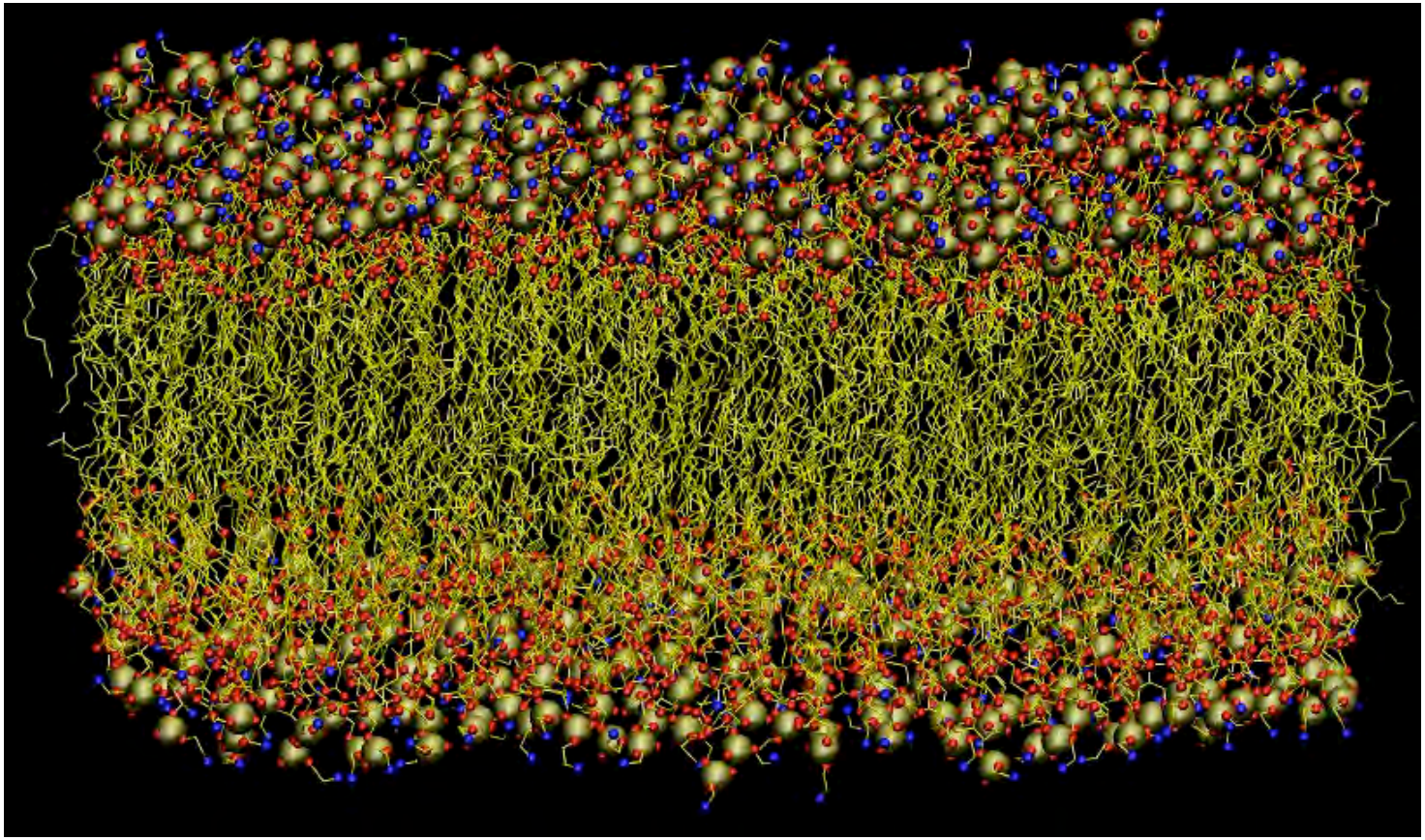
60 x 60 Å

Pure POPE

5 ns

~100,000

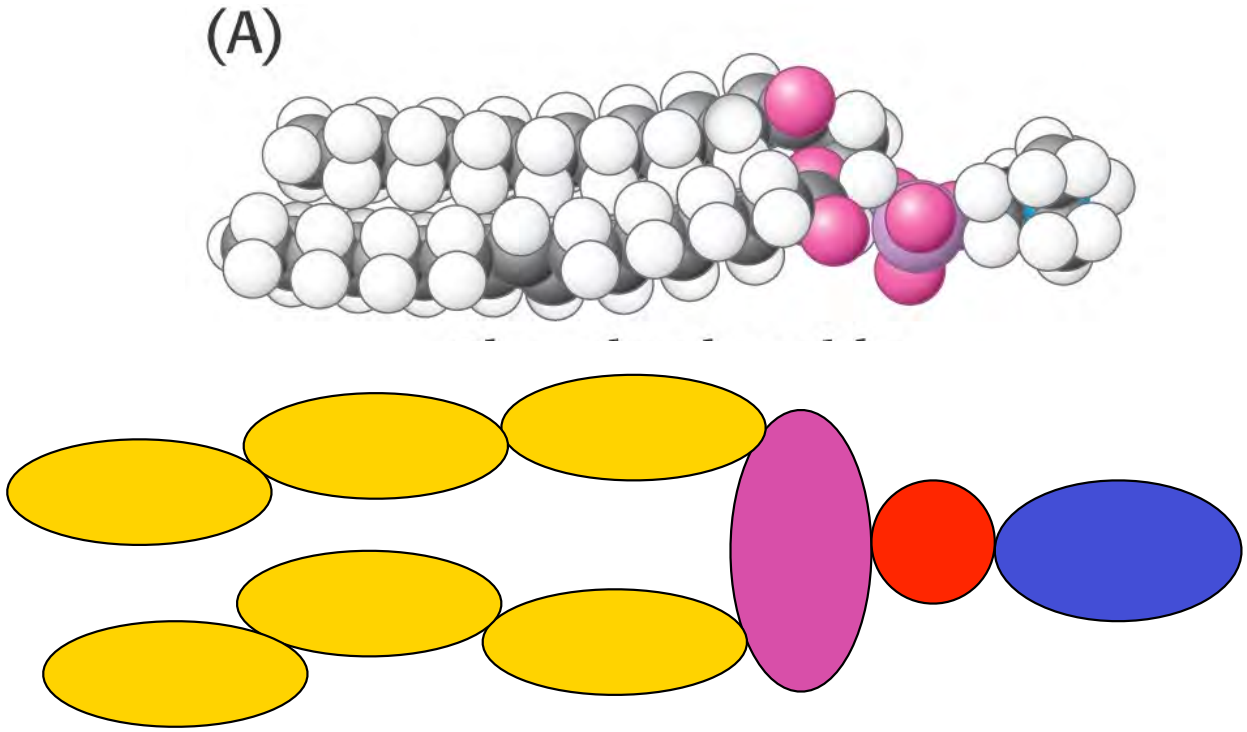
atoms



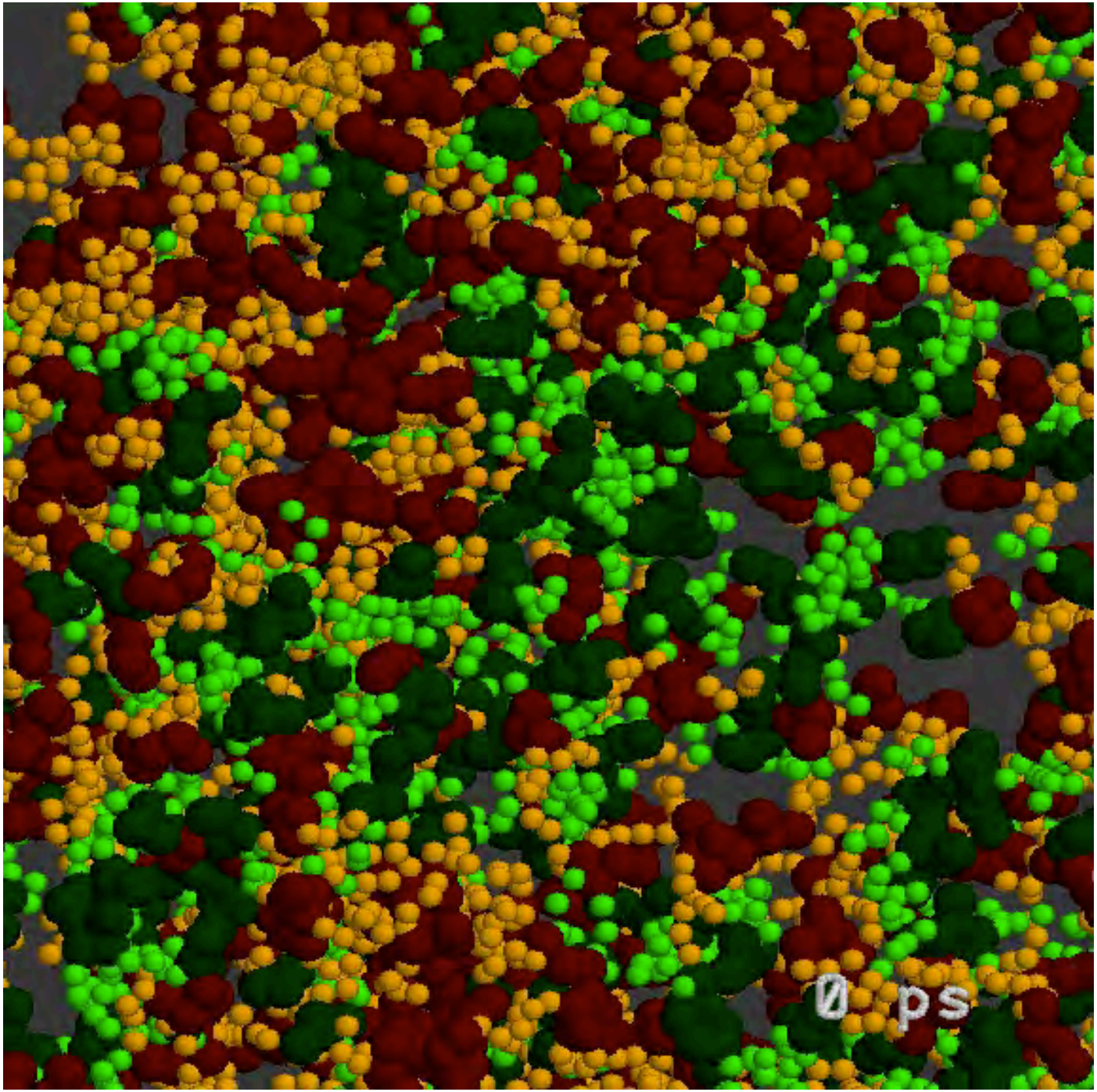
Coarse-grained modeling of lipids

150 particles

9 particles!



Also, increasing the time step by orders of magnitude.



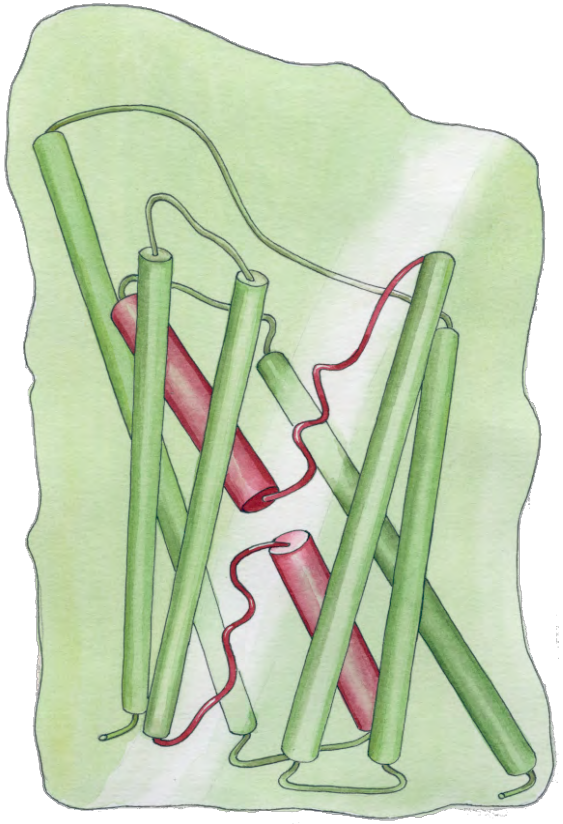
by: J. Siewert-Jan Marrink and Alan E. Mark, University of Groningen, The Netherlands

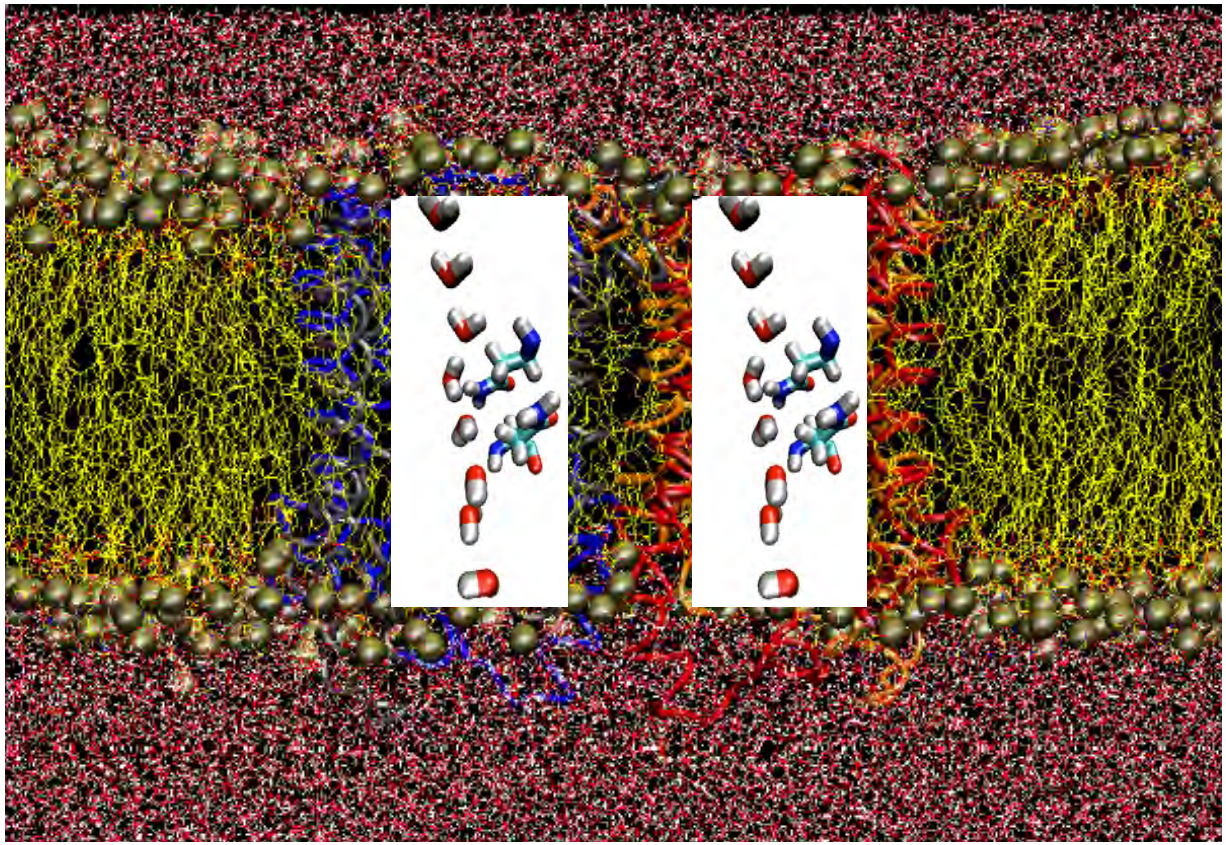
Analysis of Molecular Dynamics Simulations of Biomolecules

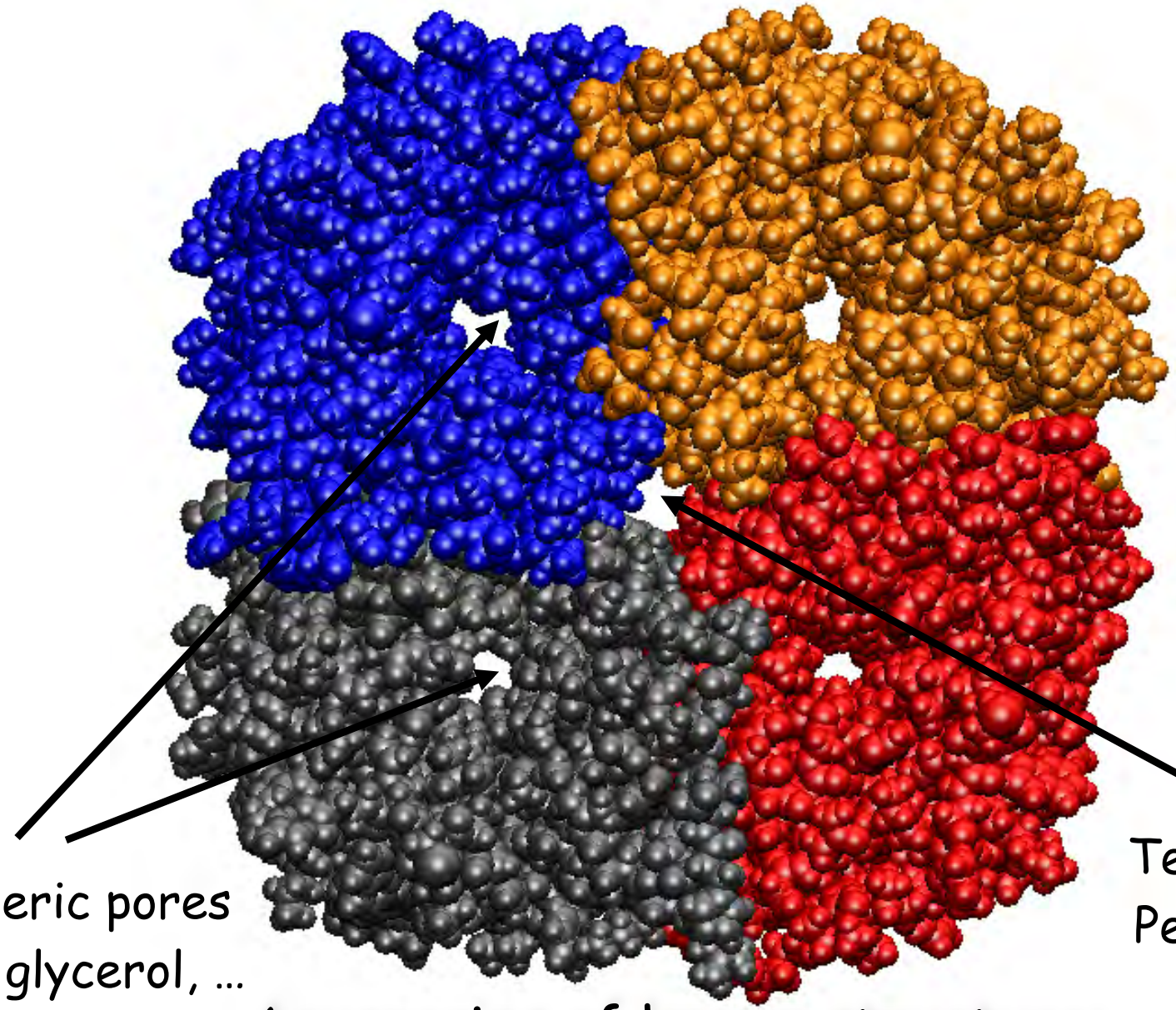
- A very complicated arrangement of hundreds of groups interacting with each other
- Where to start to look at?
- What to analyze?
- How much can we learn from simulations?

It is very important to get acquainted with your system

Aquaporins







Monomeric pores
Water, glycerol, …

Tetrameric pore
Perhaps ions???

Aquaporins of known structure:

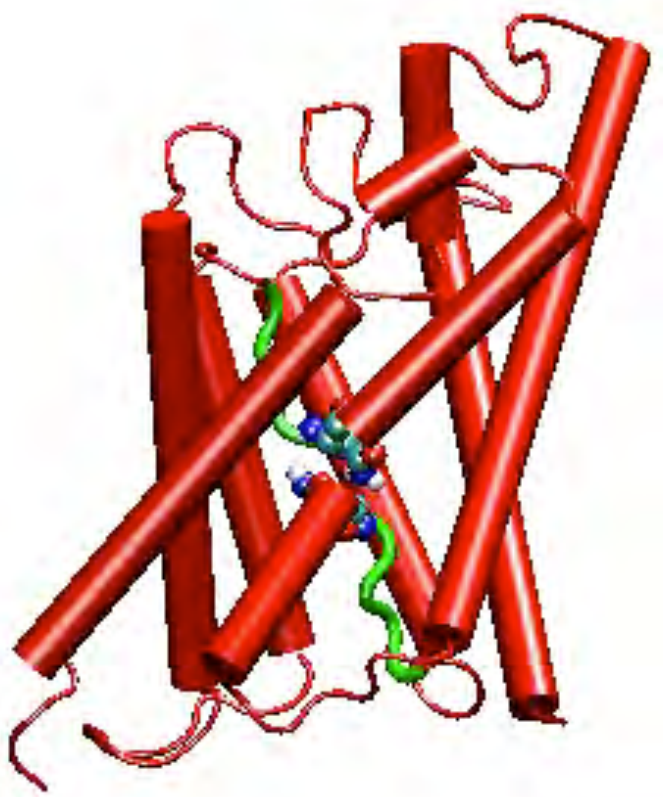
GlpF – E. coli glycerol channel (aquaglycerolporin)

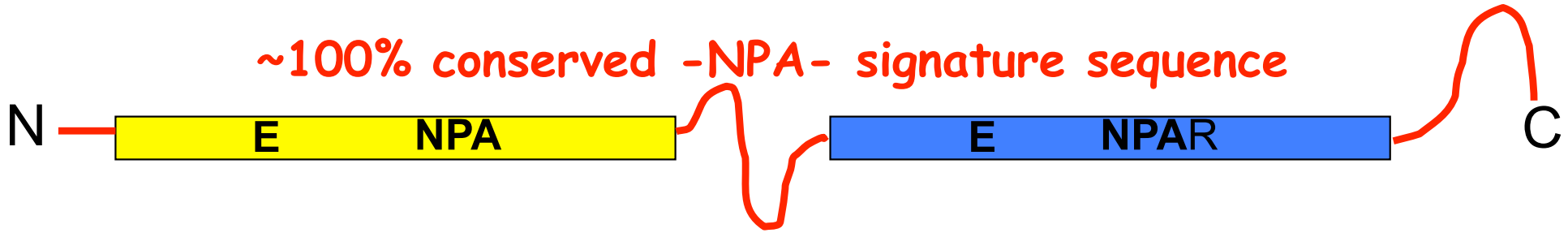
AQP1 – Mammalian aquaporin-1 (pure water channel)

AqpZ and AQP0 (2004)

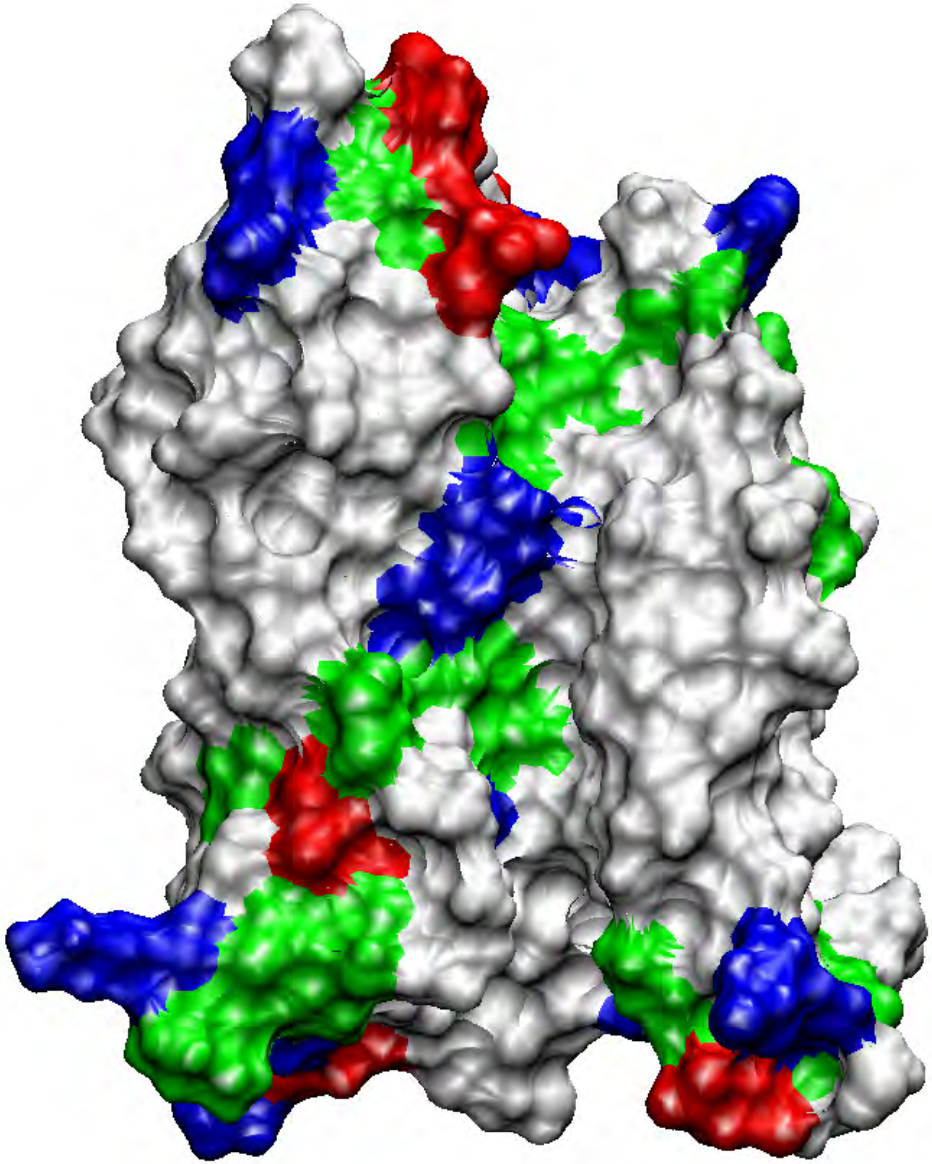
Functionally Important Features

- Tetrameric architecture
- Amphipatic channel interior
- Water and glycerol transport
- Protons, and other ions are excluded
- Conserved asparagine-proline-alanine residues; NPA motif
- Characteristic half-membrane spanning structure



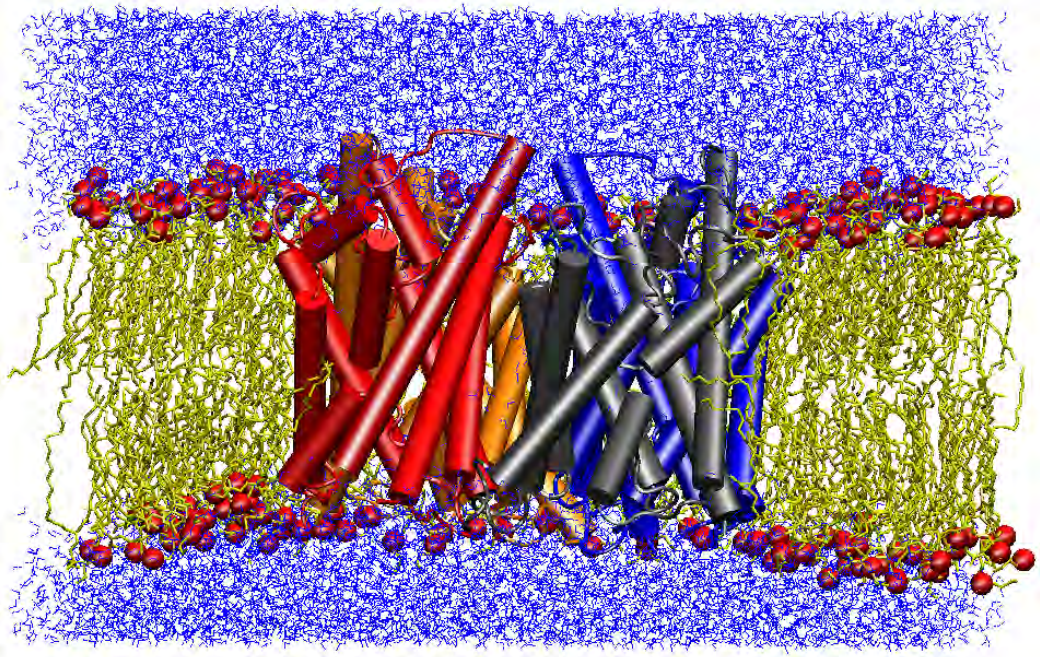


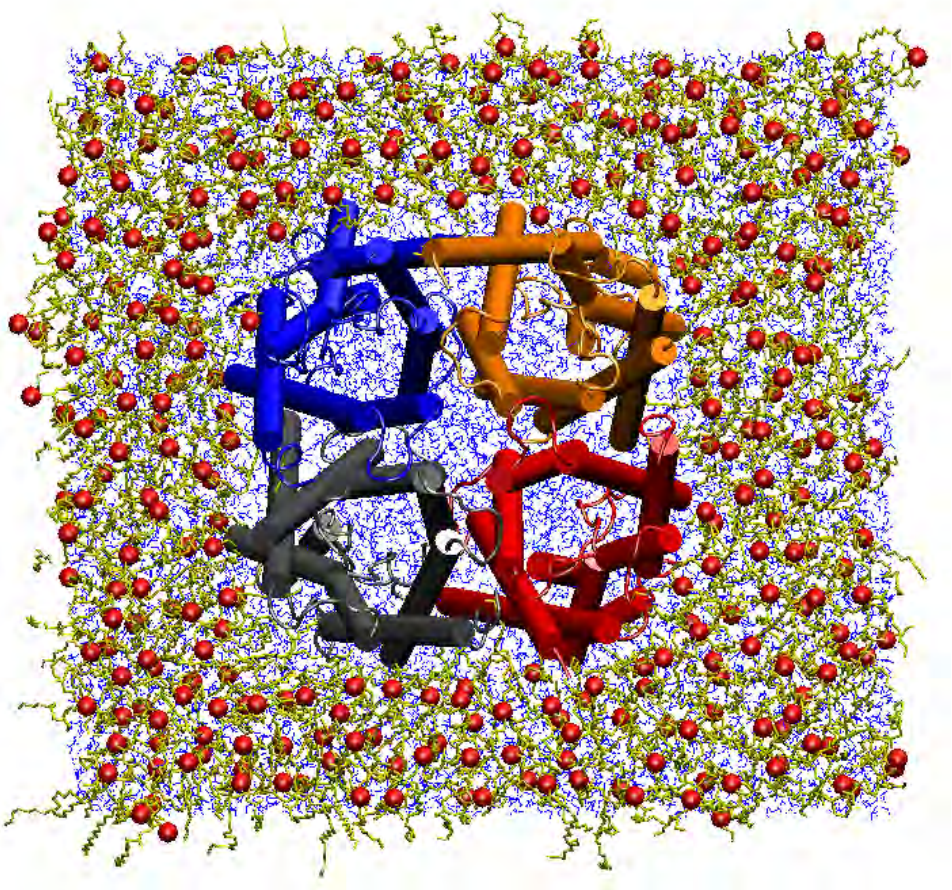
A Semi-hydrophobic channel



Molecular Dynamics Simulations

Protein:	~	15,000 atoms
Lipids (POPE):	~	40,000 atoms
Water:	~	51,000 atoms
Total:	~	106,000 atoms





NAMD, CHARMM27, PME

NpT ensemble at 310 K

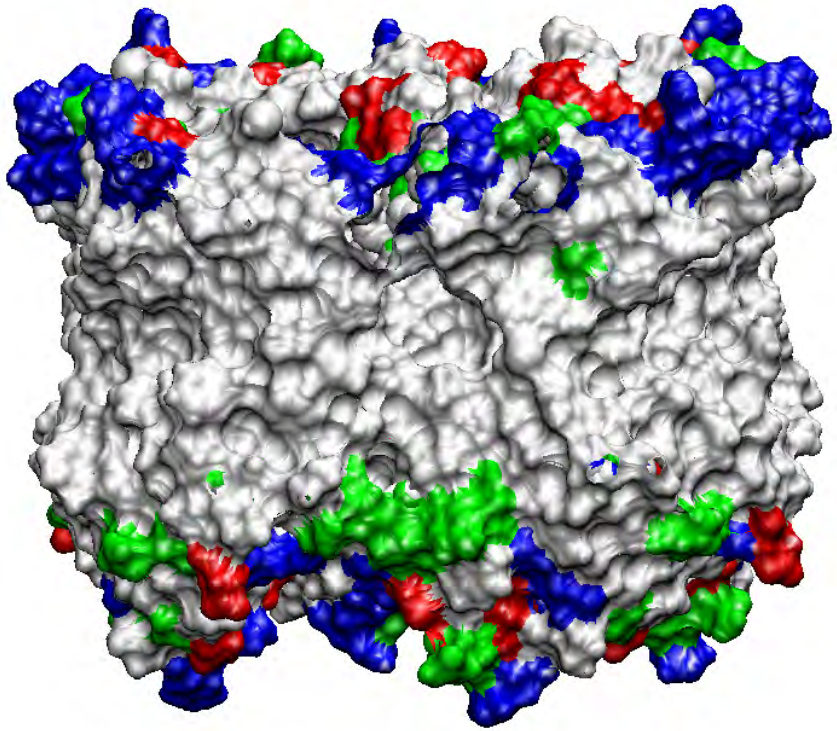
1ns equilibration, 4ns production

10 days /ns – 32-proc Linux cluster

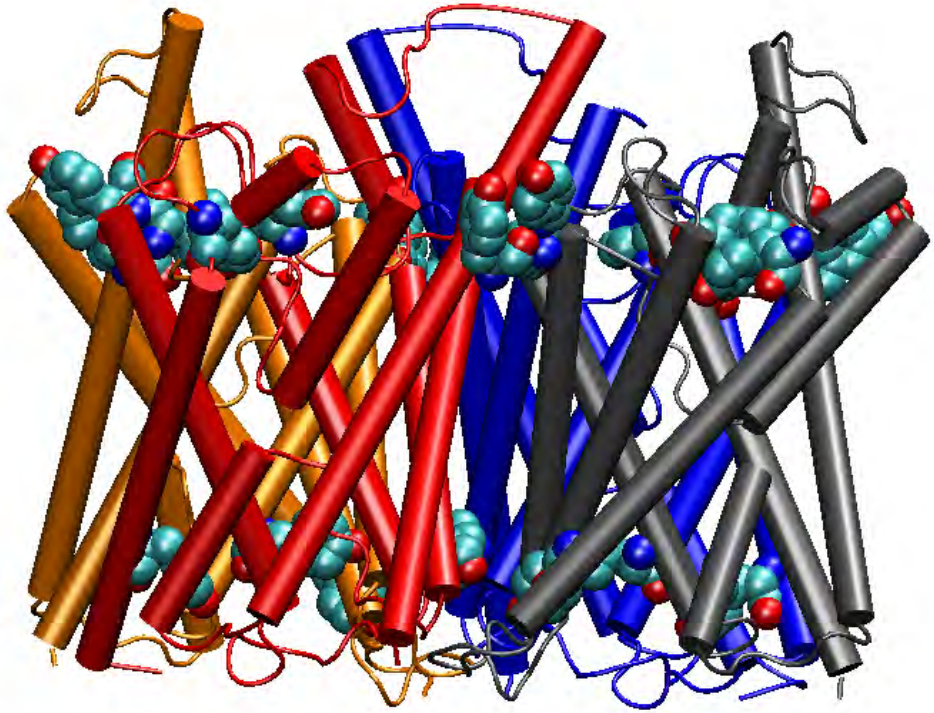
3.5 days/ns - 128 O2000 CPUs

0.35 days/ns – 512 LeMieux CPUs

Protein Embedding in Membrane



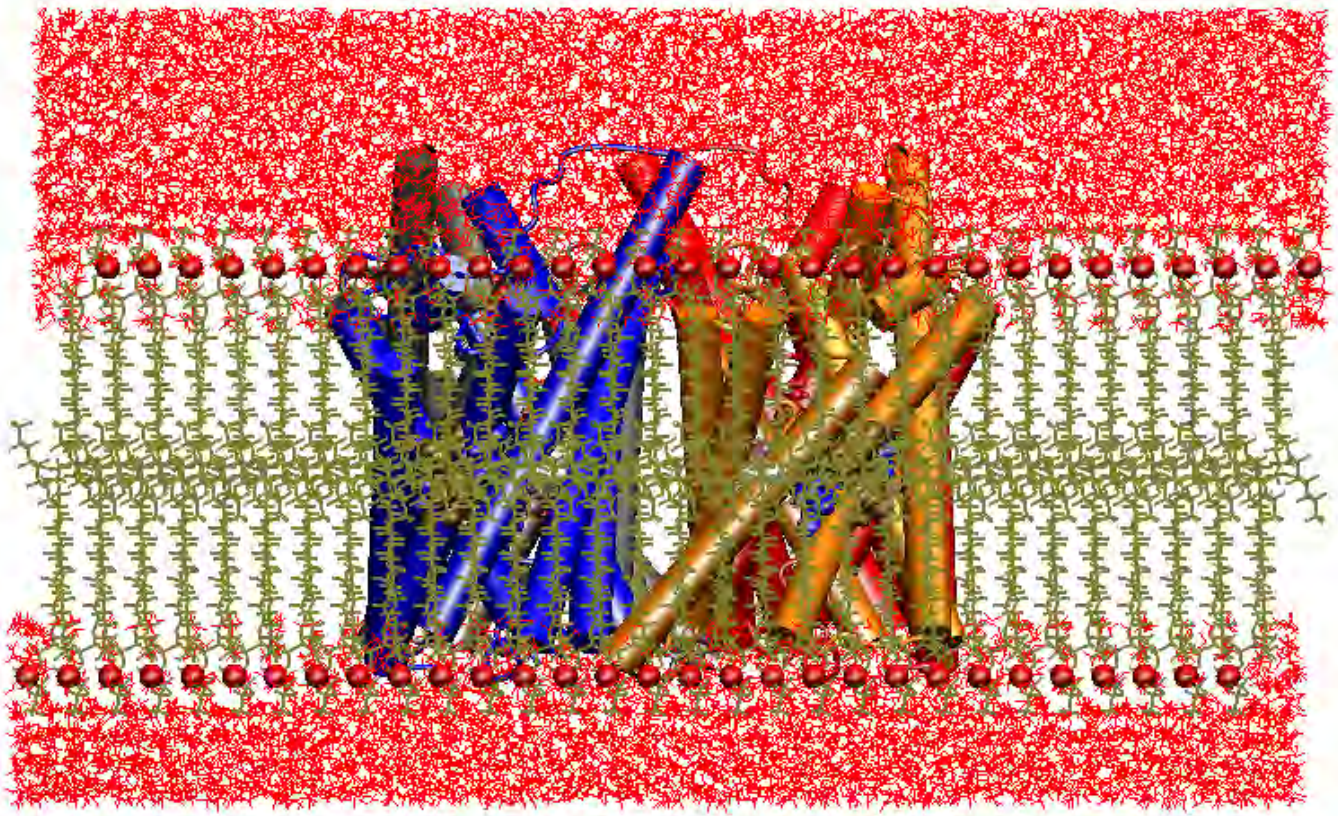
Hydrophobic surface
of the protein

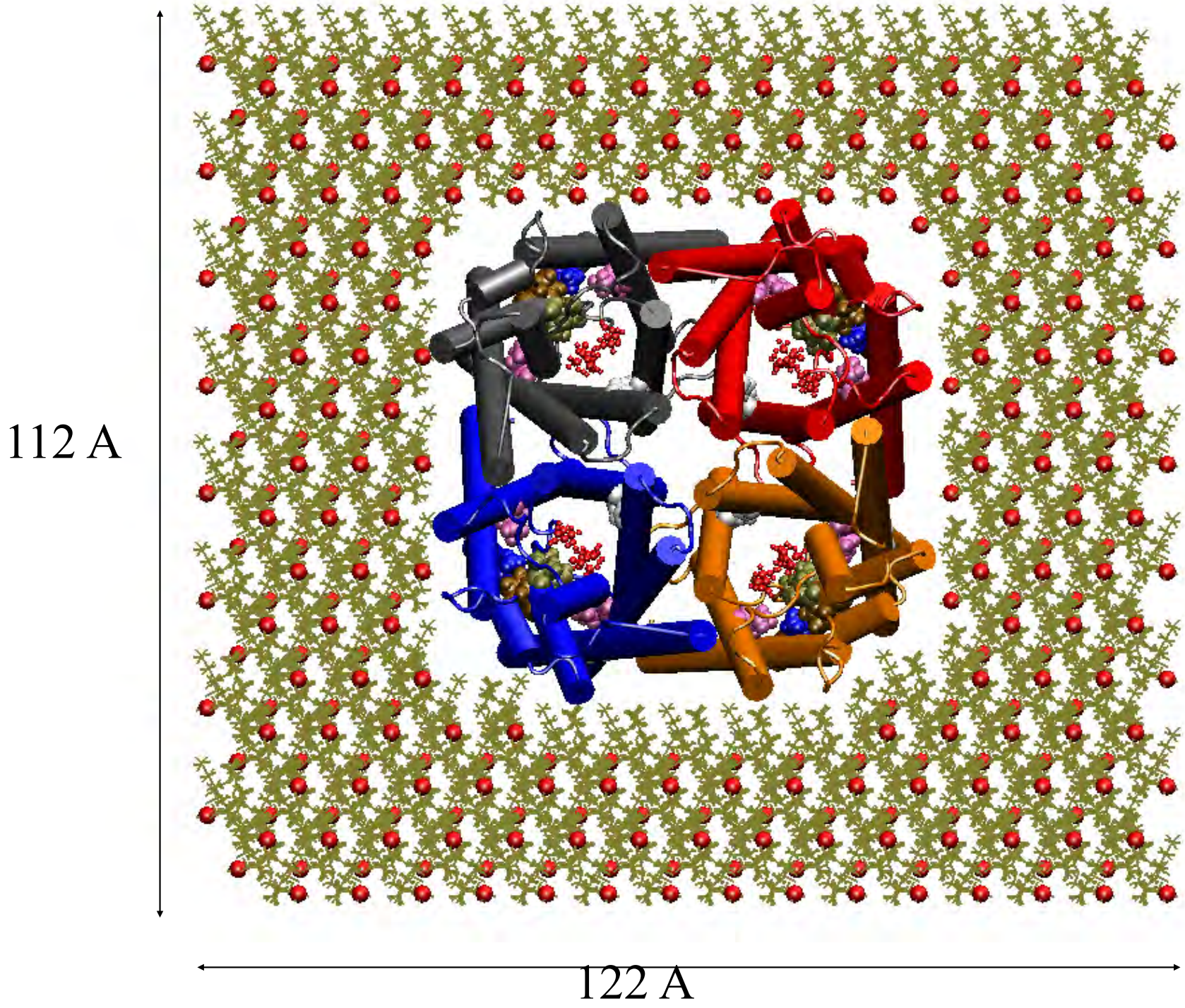


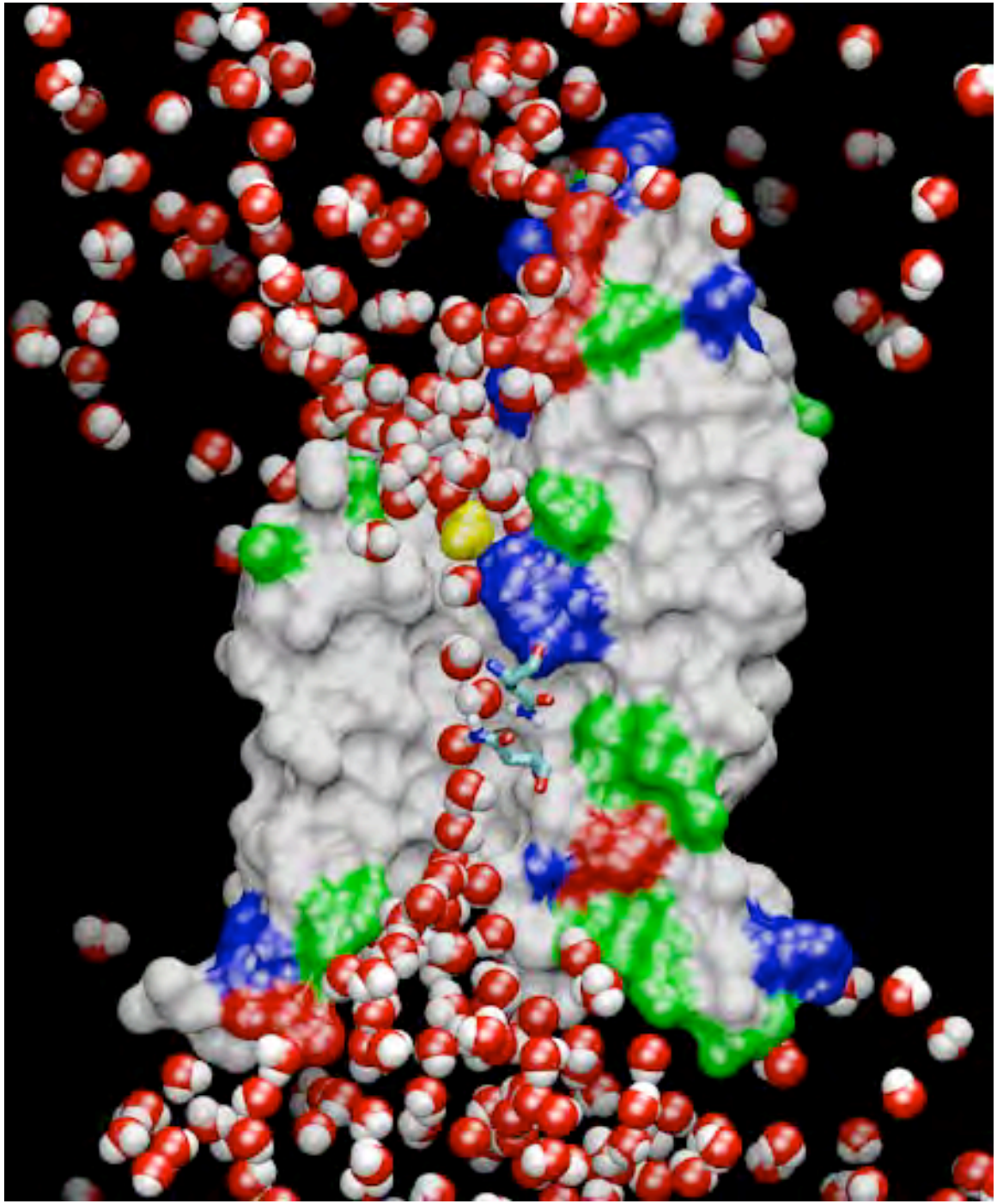
Ring of
Tyr and Trp

Embedding GlpF in Membrane

77 A

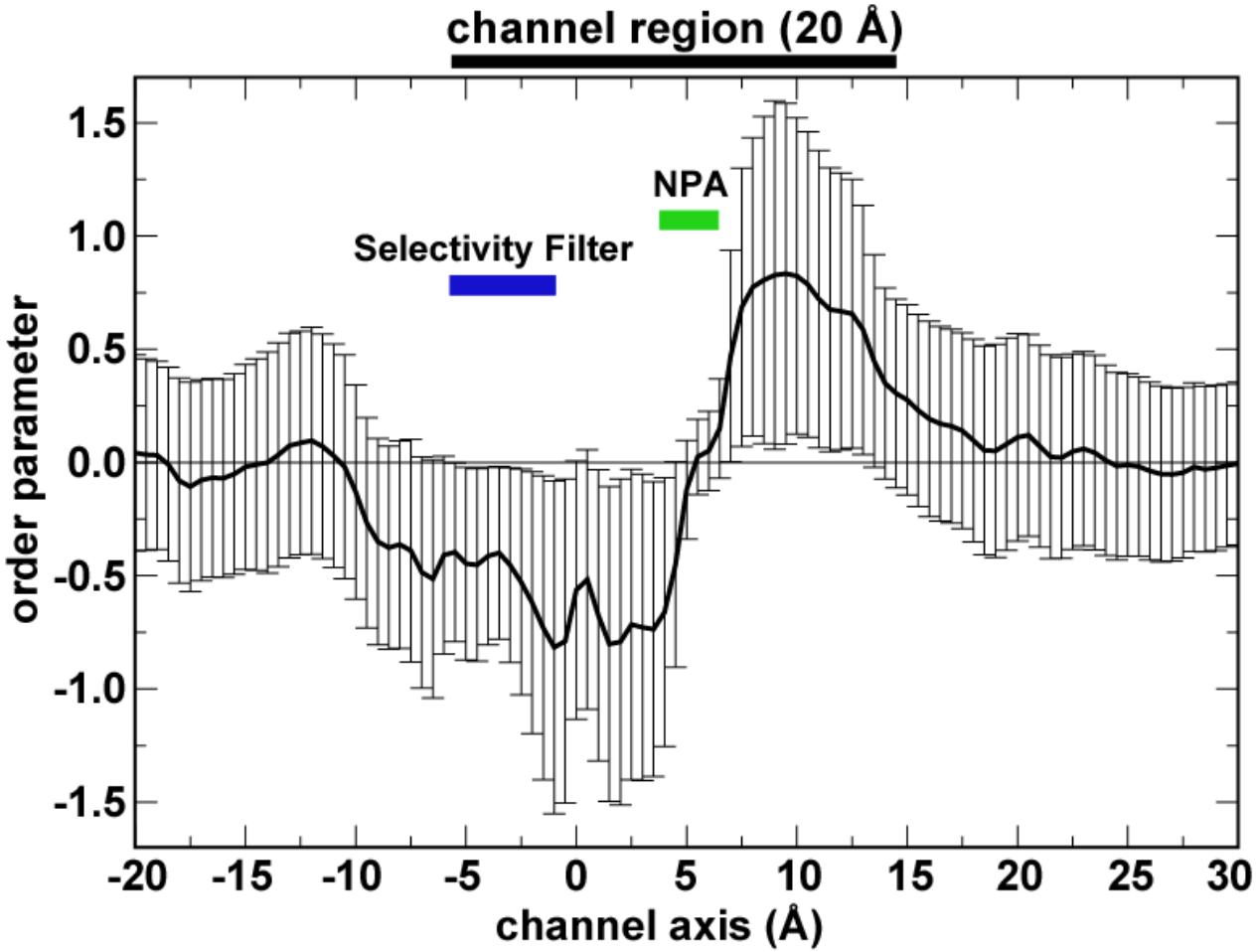



112 A
122 A



Animation available at the Nobel web site

E. T., et al., ***Science*** 2002.

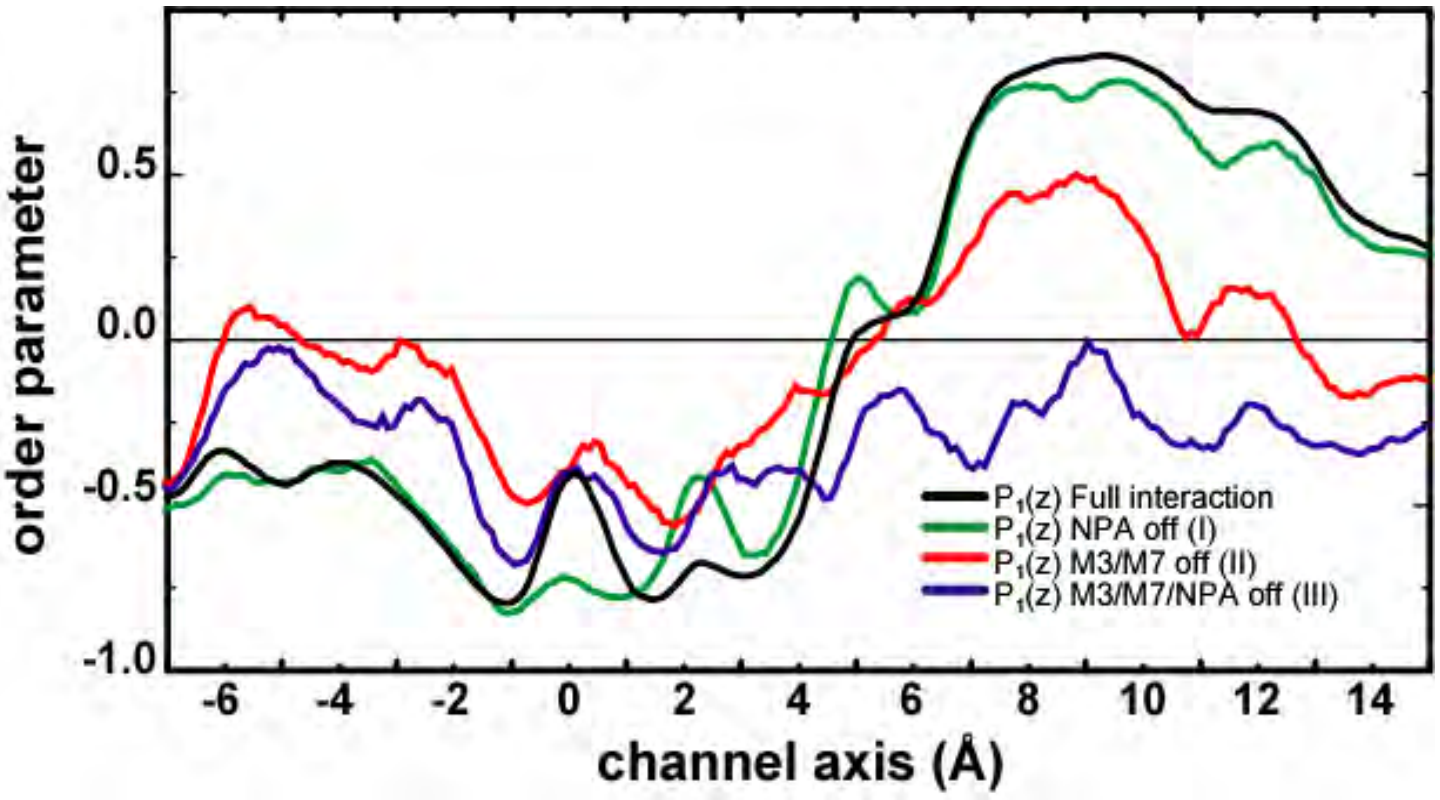

channel region (20 Å)
NPA
Selectivity Filter
order parameter
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-20
-15
-10
-5
0
5
10
15
20
25
30
channel axis (Å)

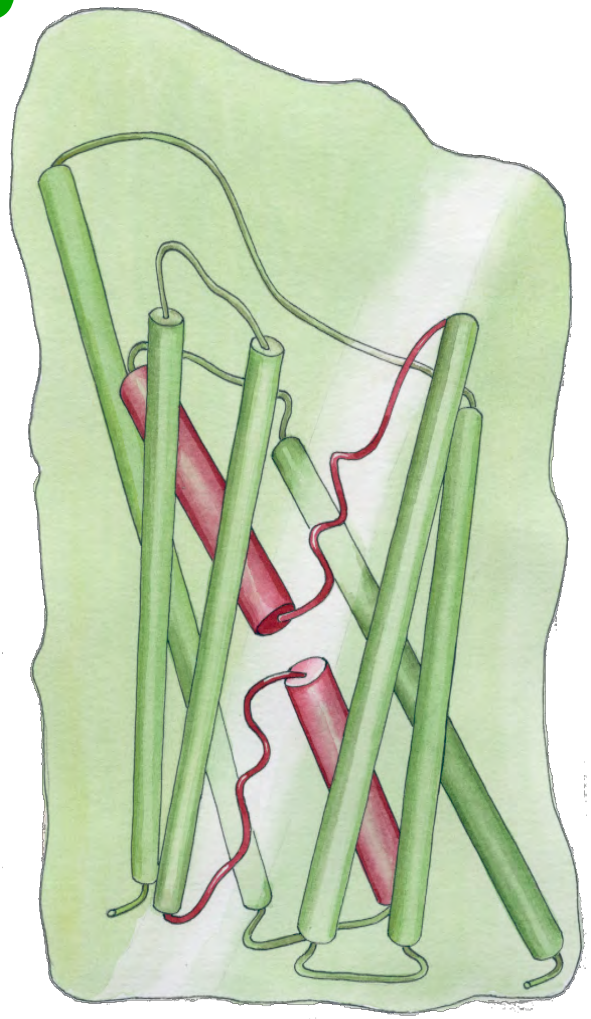
R E M E M B E R:

One of the most useful advantages of simulations over experiments is that you can modify the system as you wish: You can do modifications that are not even possible at all in reality!

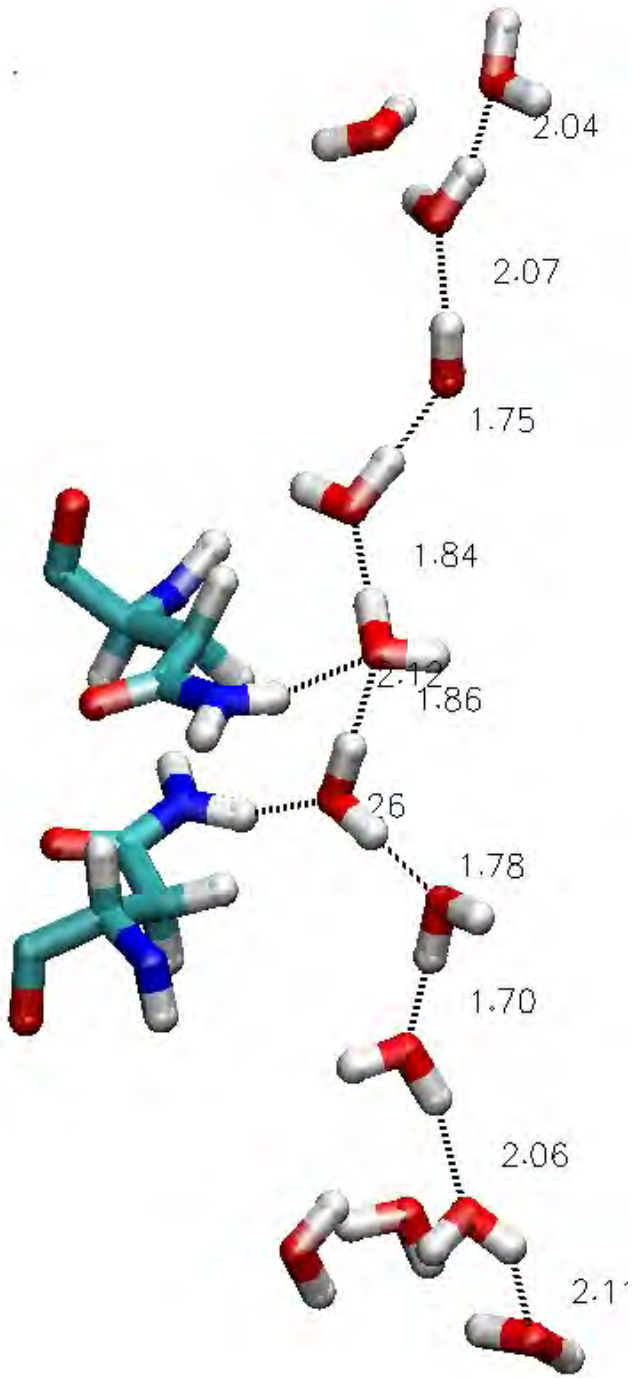
This is a powerful technique to test hypotheses developed during your simulations. Use it!

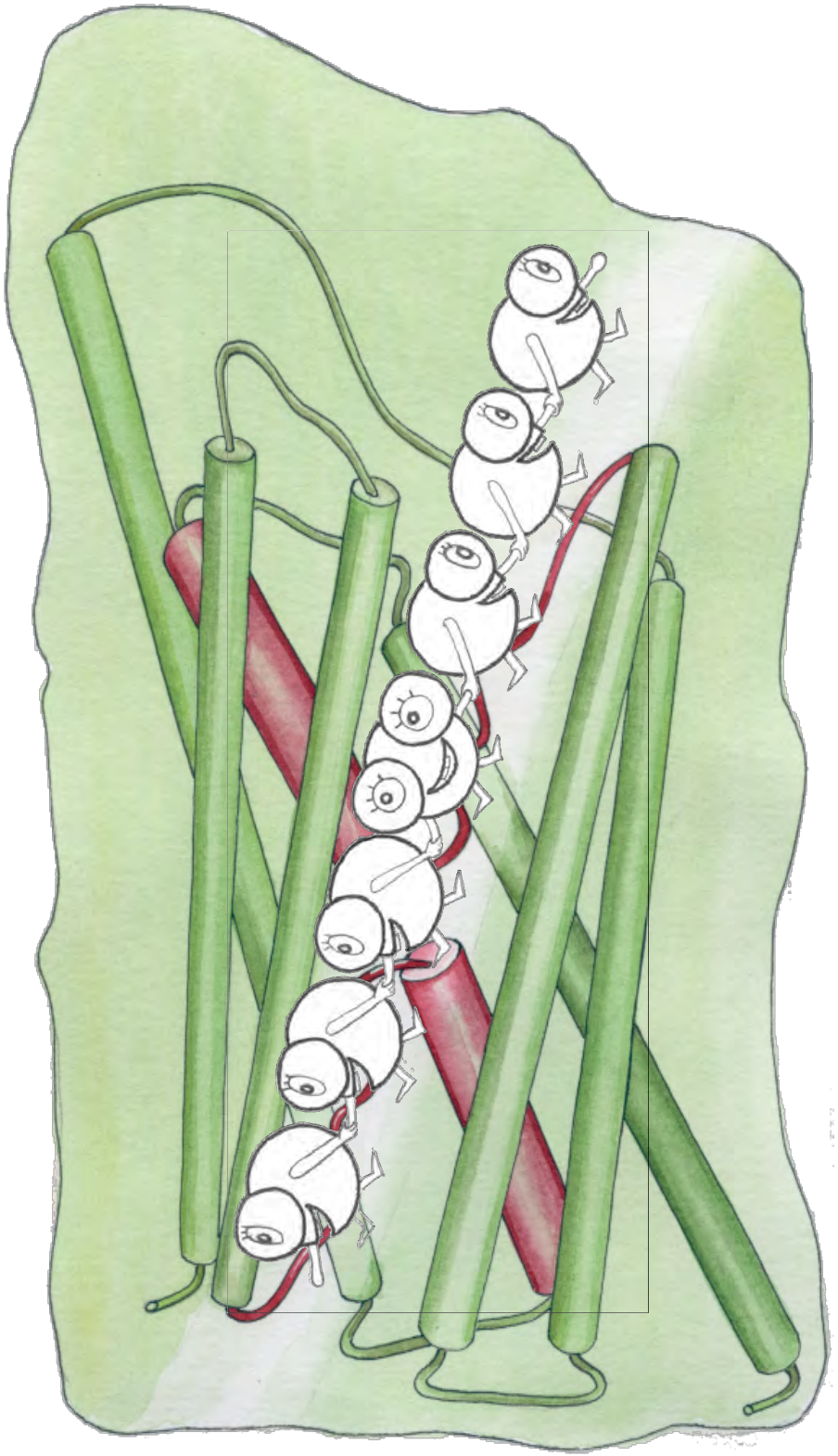
Electrostatic Stabilization of Water Bipolar Arrangement



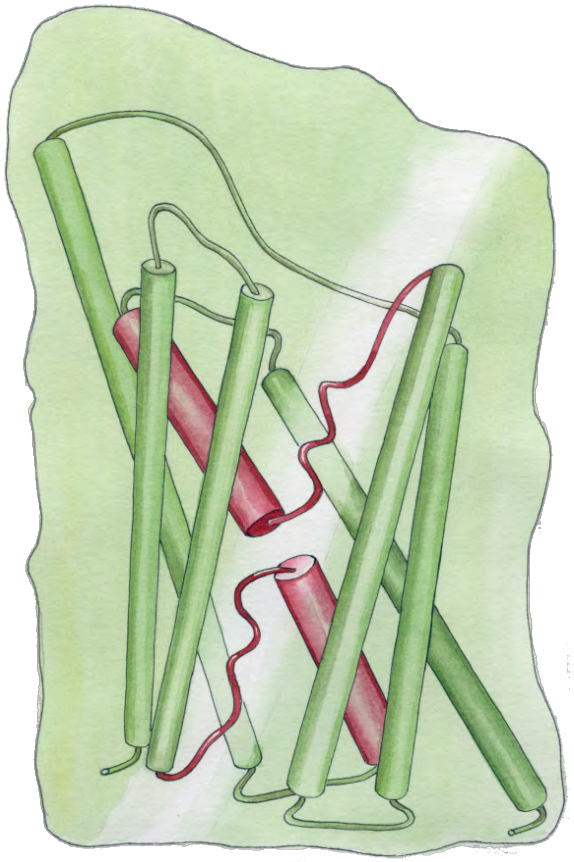


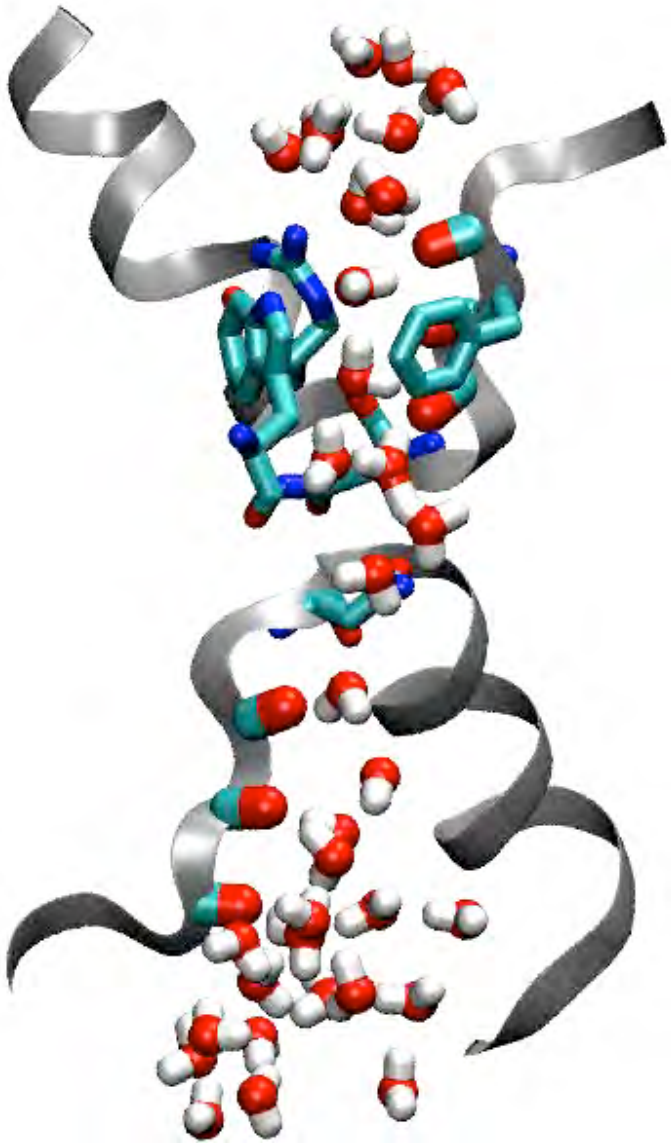
E. T., et al., ***Science*** 2002.


2.04
2.07
1.75
1.84
2.12
1.86
26
1.78
1.70
2.06
2.1



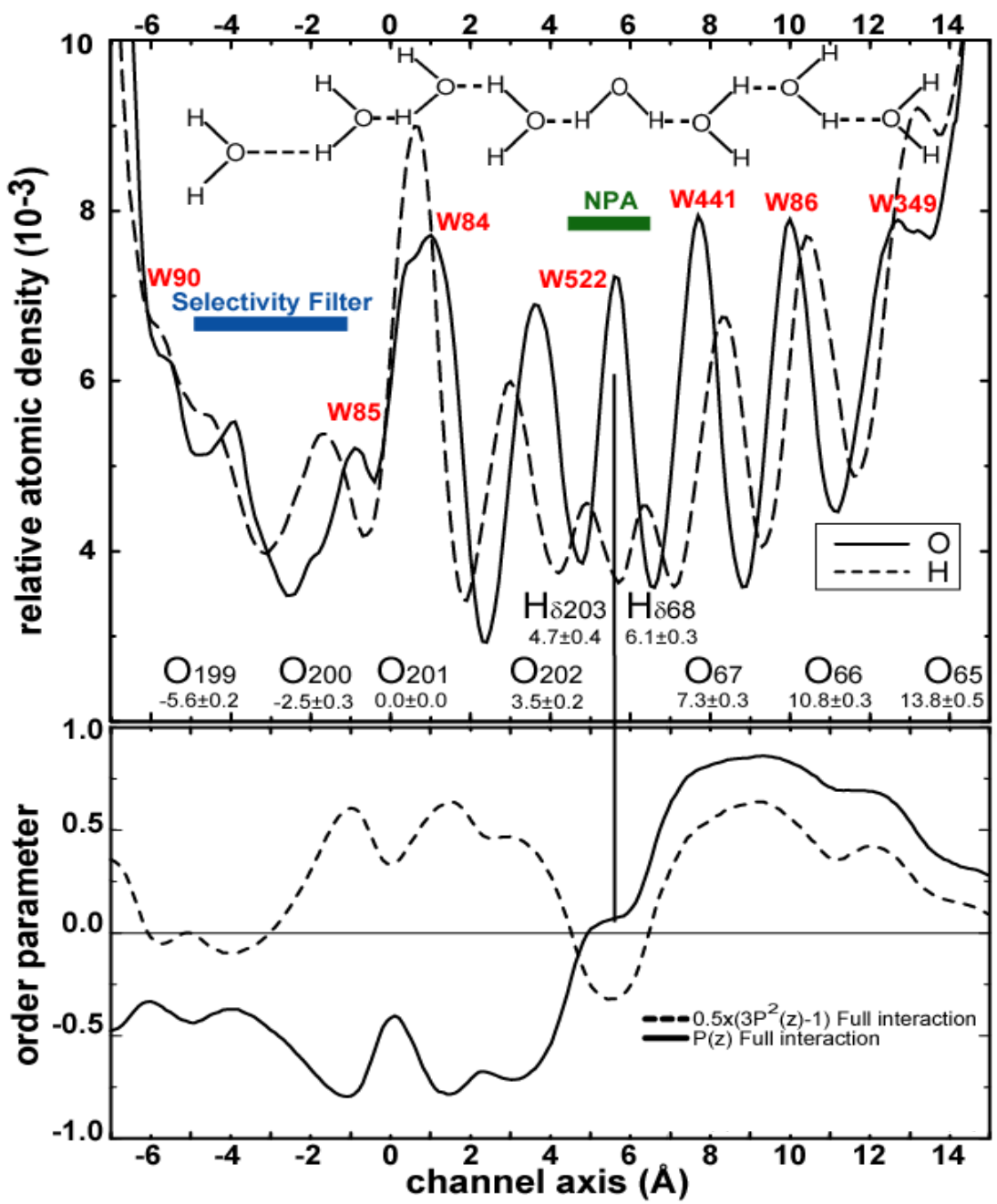
Characterizing Protein Forces

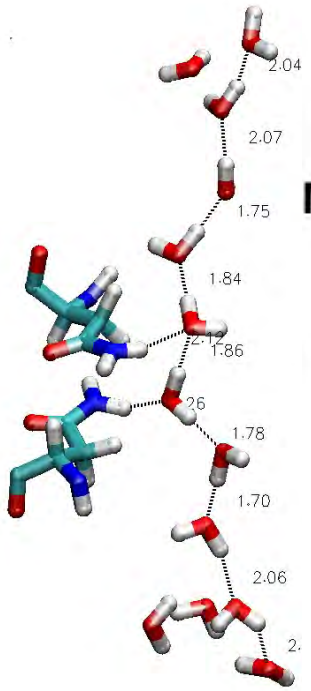


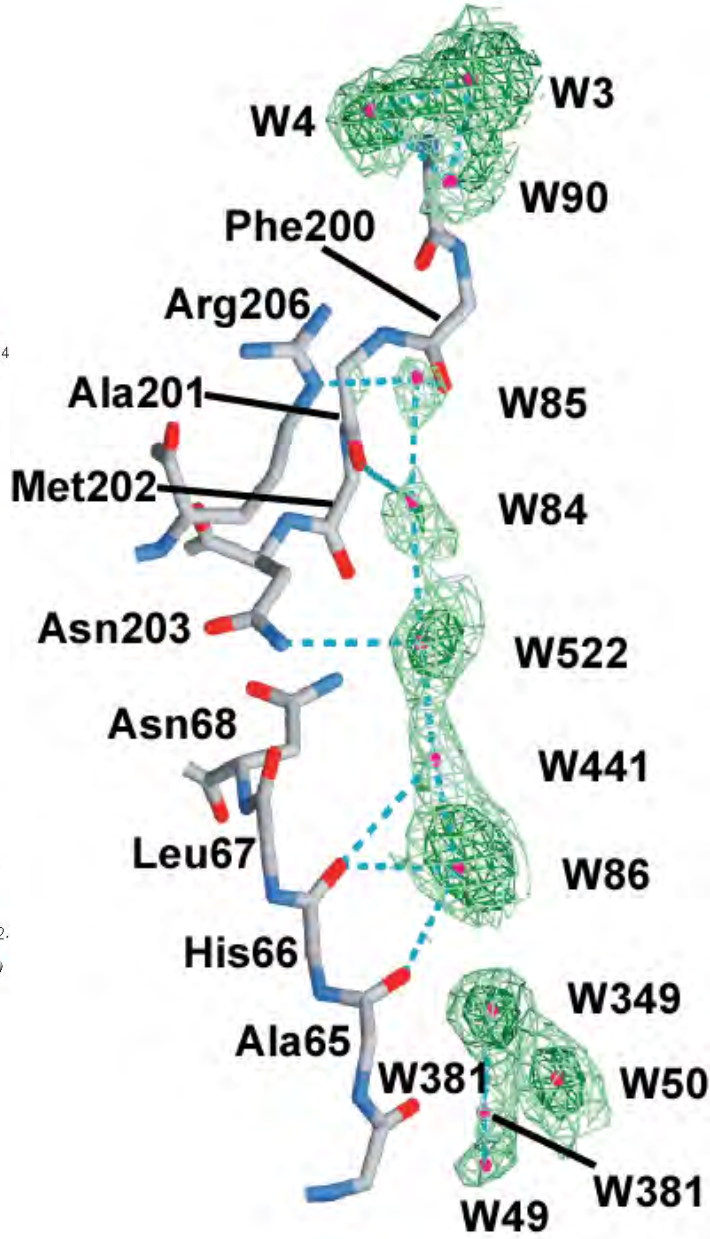


QM/MM MD of proton behavior in the channel

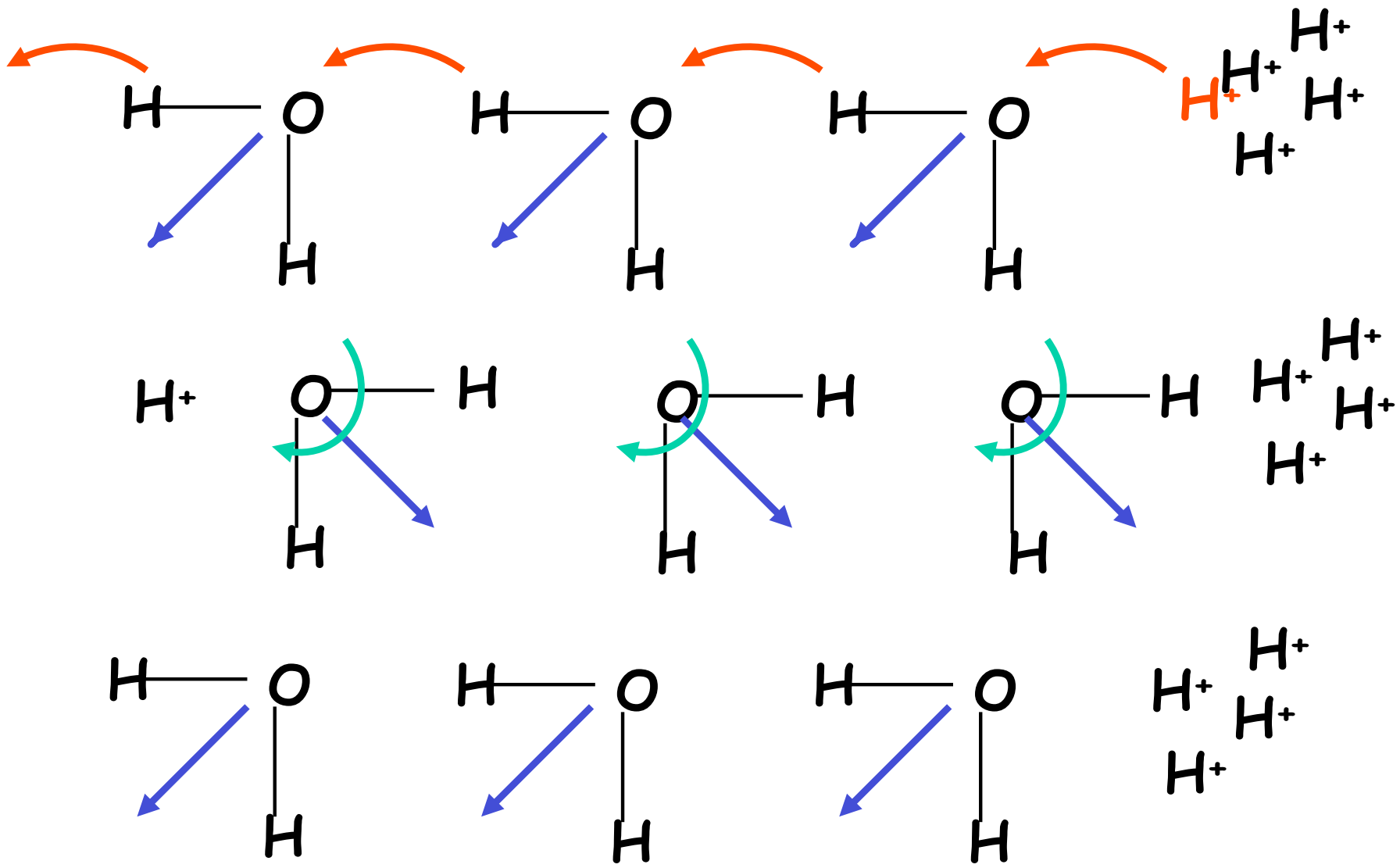
Water Bipolar Configuration in Aquaporins







Proton transfer through water



Battling the Timescale - Case I

Steered Molecular Dynamics is a non-equilibrium method by nature

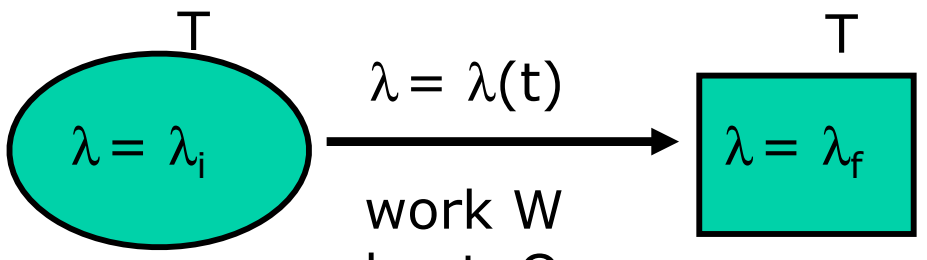
- A wide variety of events that are inaccessible to conventional molecular dynamics simulations can be probed.

- The system will be driven, however, away from equilibrium, resulting in problems in describing the energy landscape associated with the event of interest.

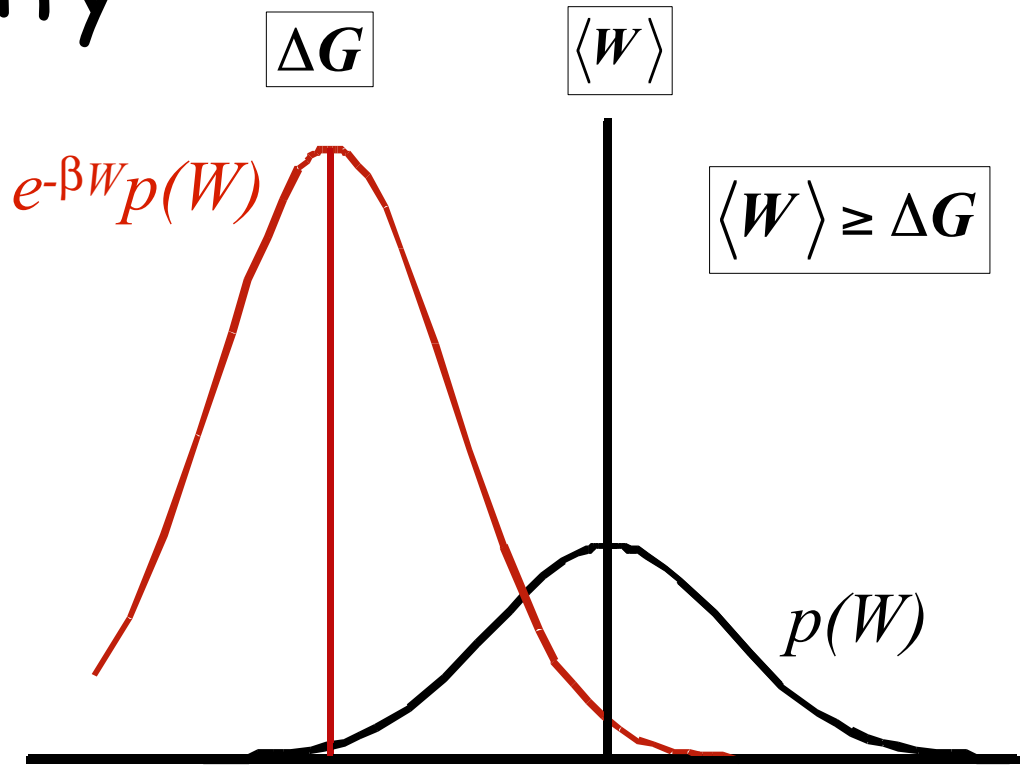
Second law of thermodynamics $\longrightarrow W \geq \Delta G$

Jarzynski's Equality

Transition between two equilibrium states



$$\Delta G = G_f - G_i$$



C. Jarzynski, *Phys. Rev. Lett.*, **78**, 2690 (1997)
C. Jarzynski, *Phys. Rev. E,* **56**, 5018 (1997)

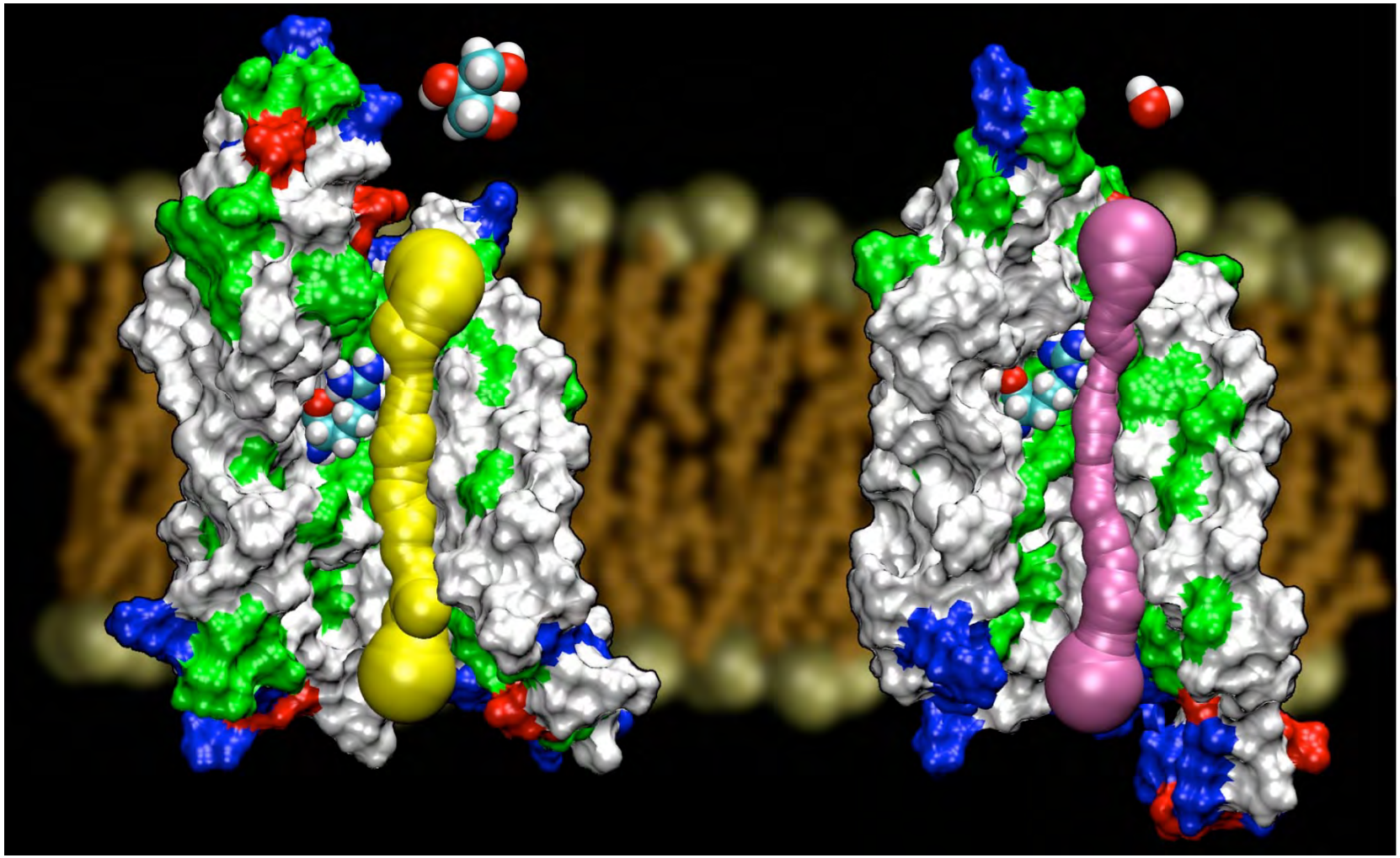
$$\left\langle e^{-\beta W} \right\rangle = e^{-\beta \Delta G}$$

$$\beta = 1/k_B T$$

In principle, it is possible to obtain free energy surfaces from repeated non-equilibrium experiments.

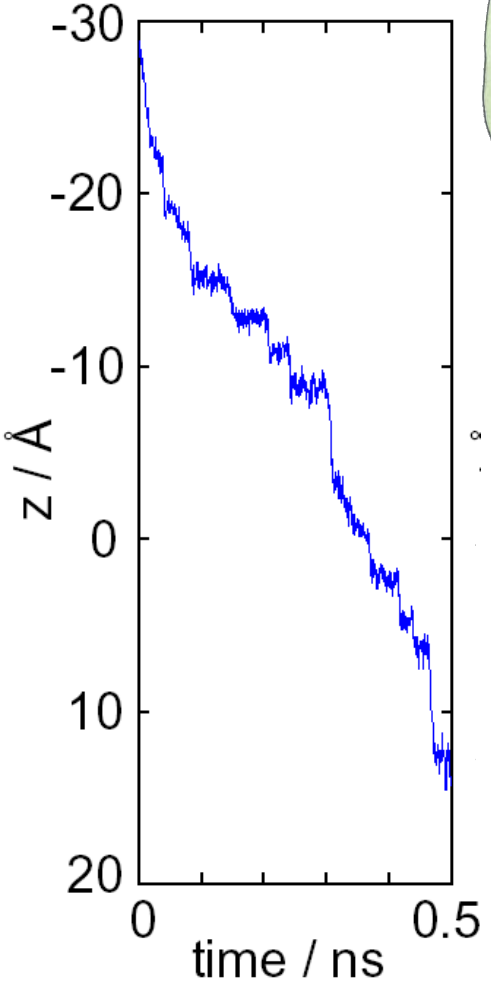
AqpZ *vs*. GlpF

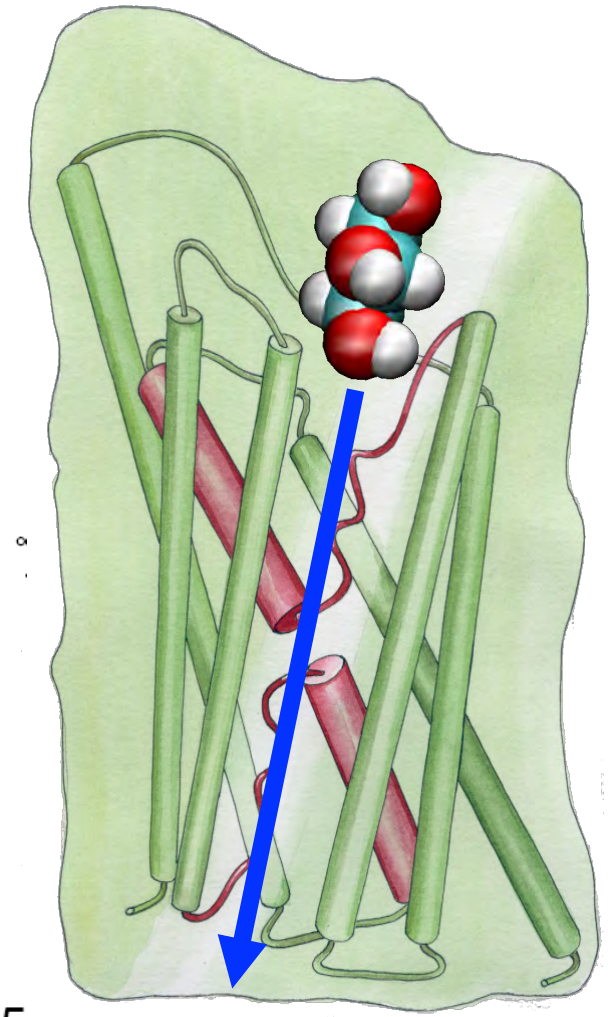
- Both from *E. coli*
- AqpZ is a pure water channel
- GlpF is a glycerol channel
- We have high resolution structures for both channels



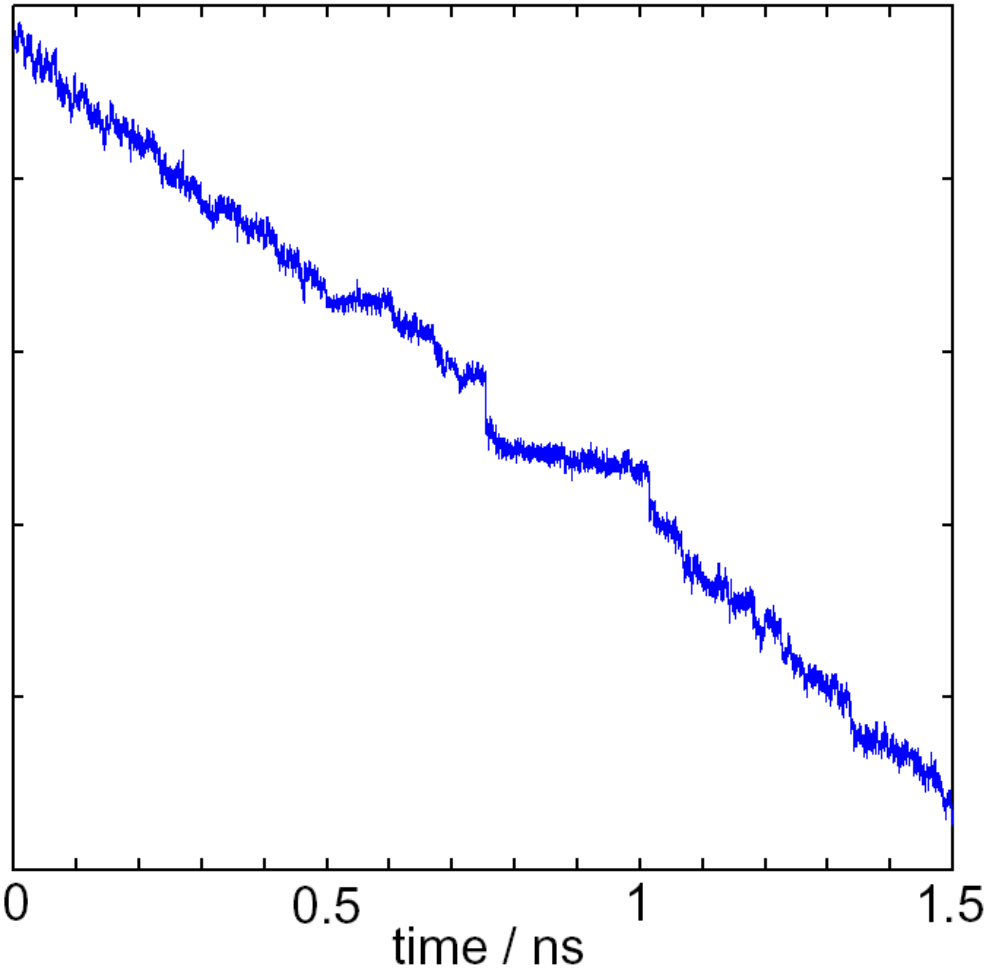
Steered Molecular Dynamics

constant force
(250 pN)

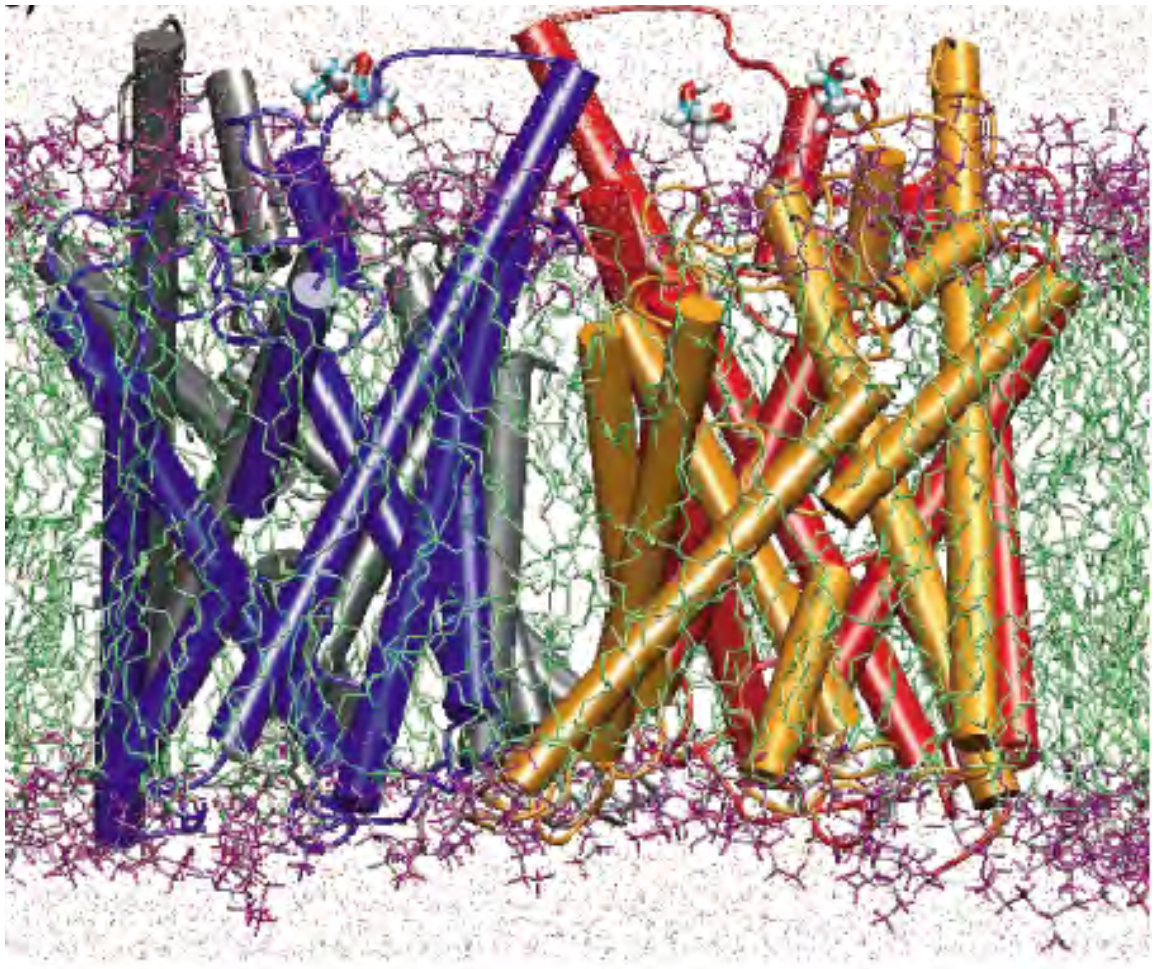


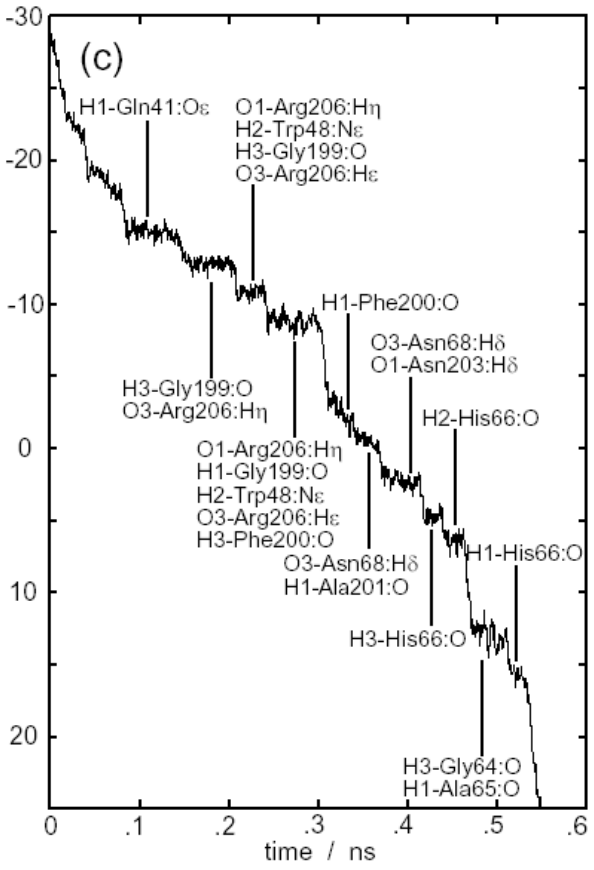


constant velocity
(30 Å/ns)



SMD Simulation of Glycerol Passage





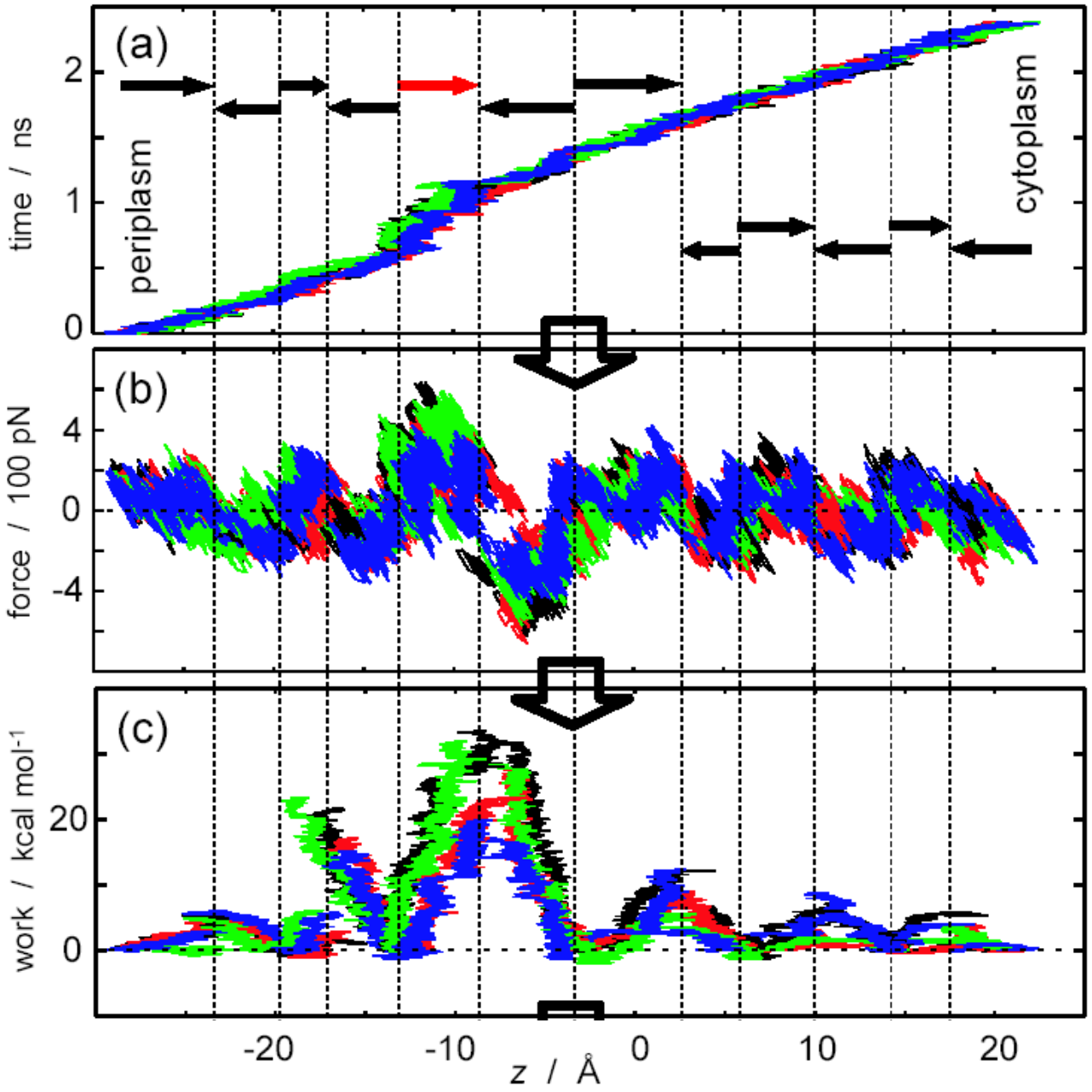
Trajectory of glycerol pulled by **constant force**

Constructing the Potential of Mean Force

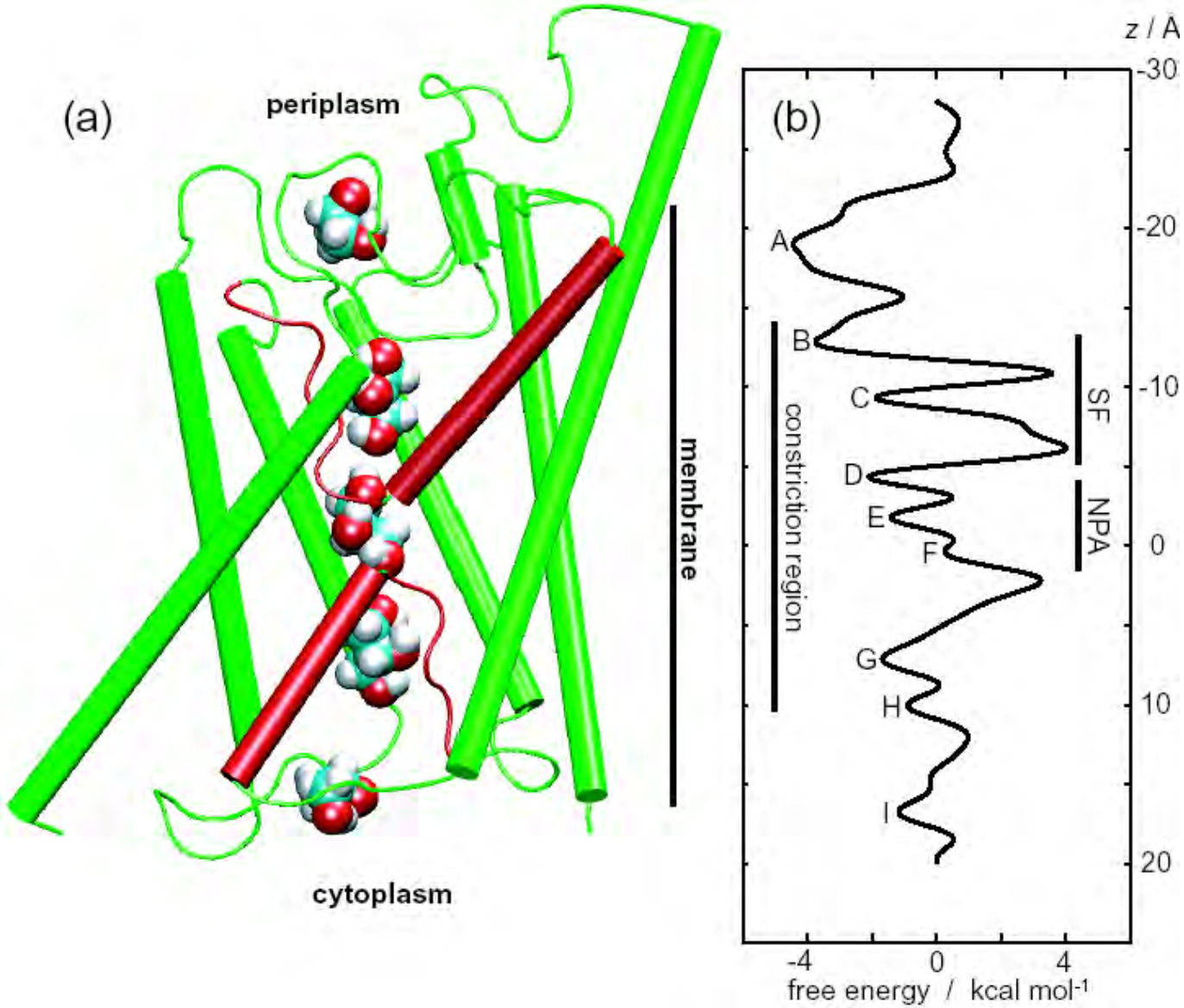
4 trajectories
v = 0.03, 0.015 Å/ps
k = 150 pN/Å

$$f(t) = -k[z(t) - z_0 - vt]$$

$$W(t) = \int_0^t dt'\, vf(t')$$



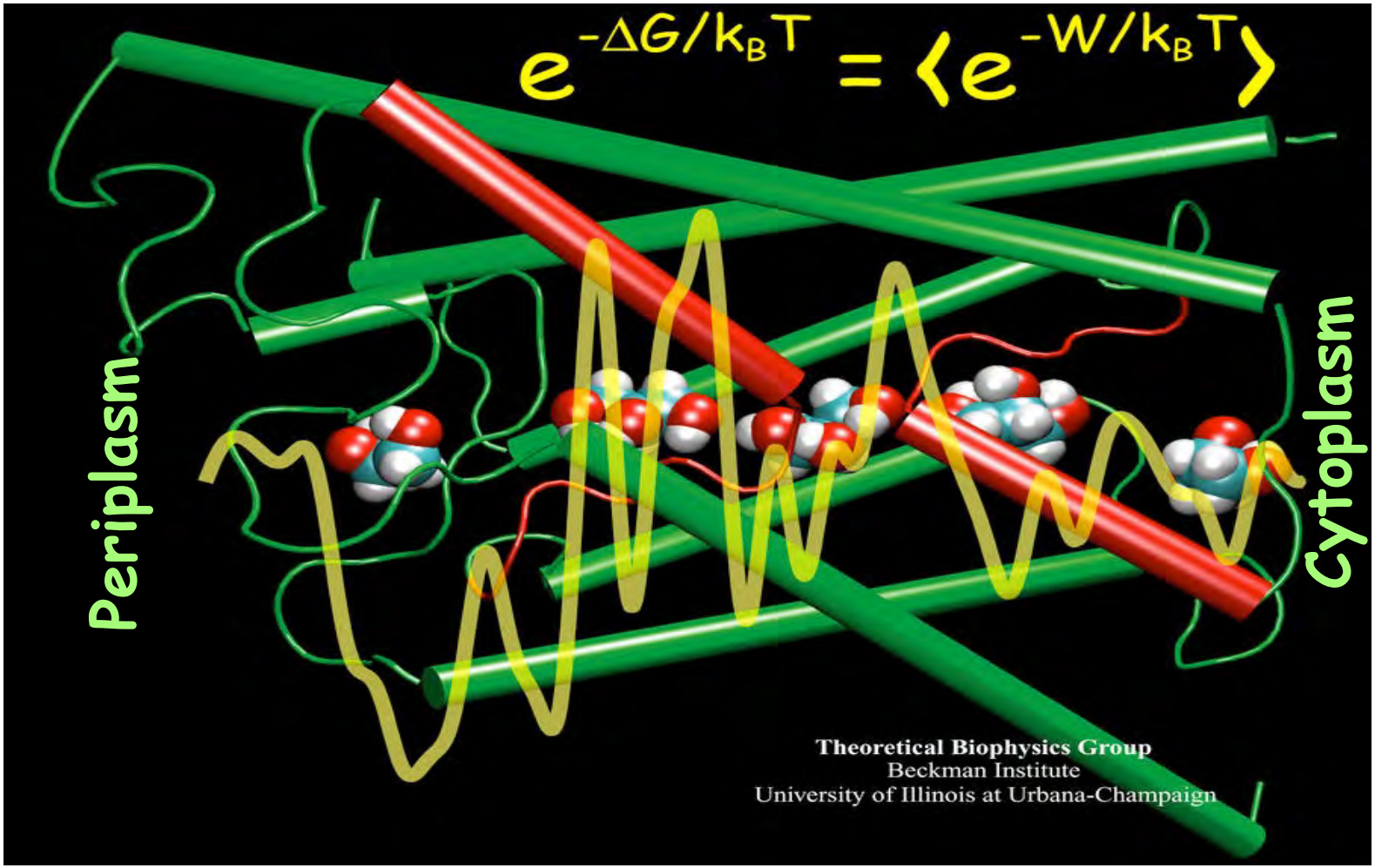
Features of the Potential of Mean Force



- Captures major features of the channel
- The largest barrier ≈**7.3 kcal/mol; exp.: 9.6±1.5 kcal/mol**

Jensen et al., *PNAS*, 99:6731-6736, 2002.

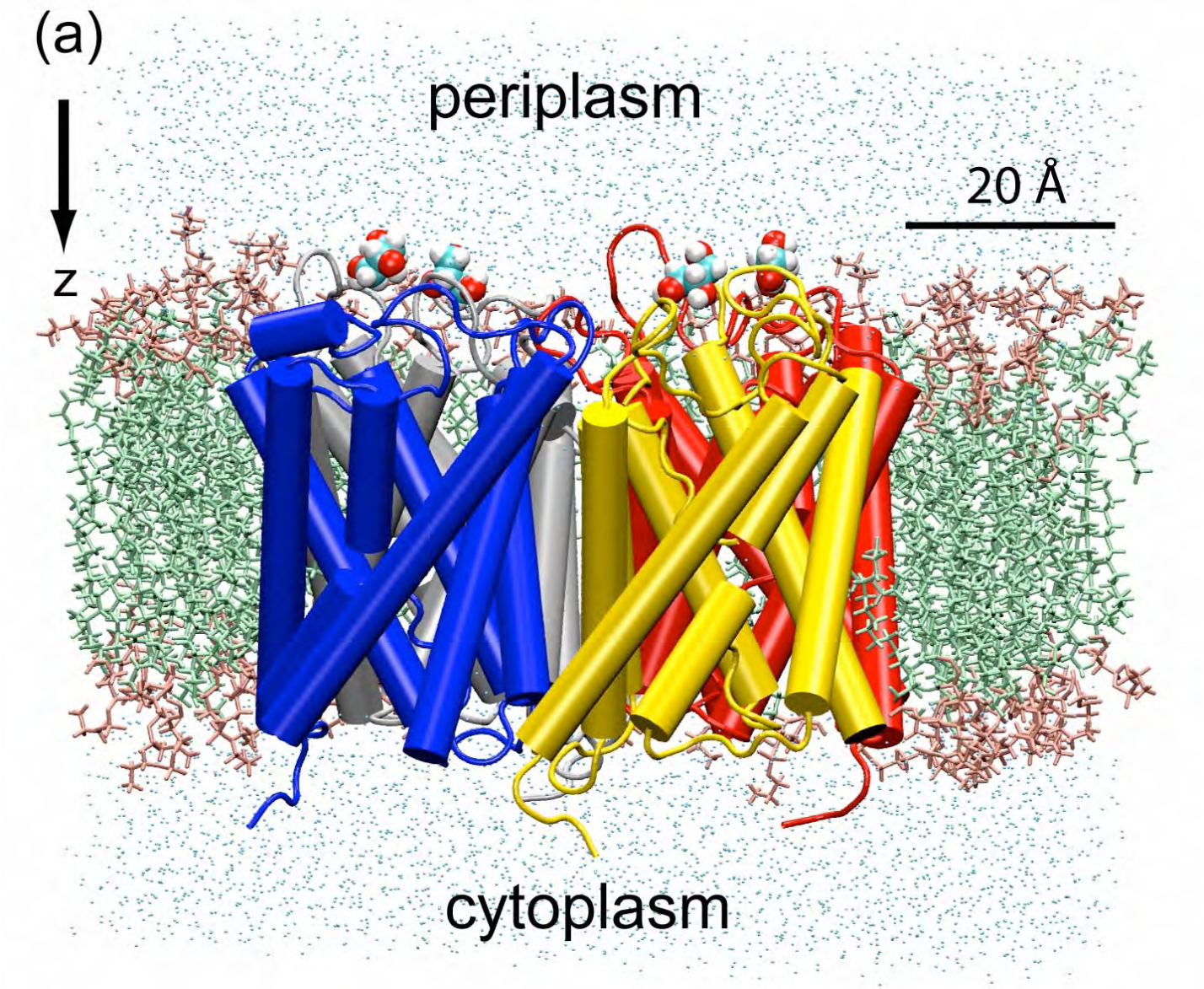
Features of the Potential of Mean Force



Asymmetric Profile in the Vestibules

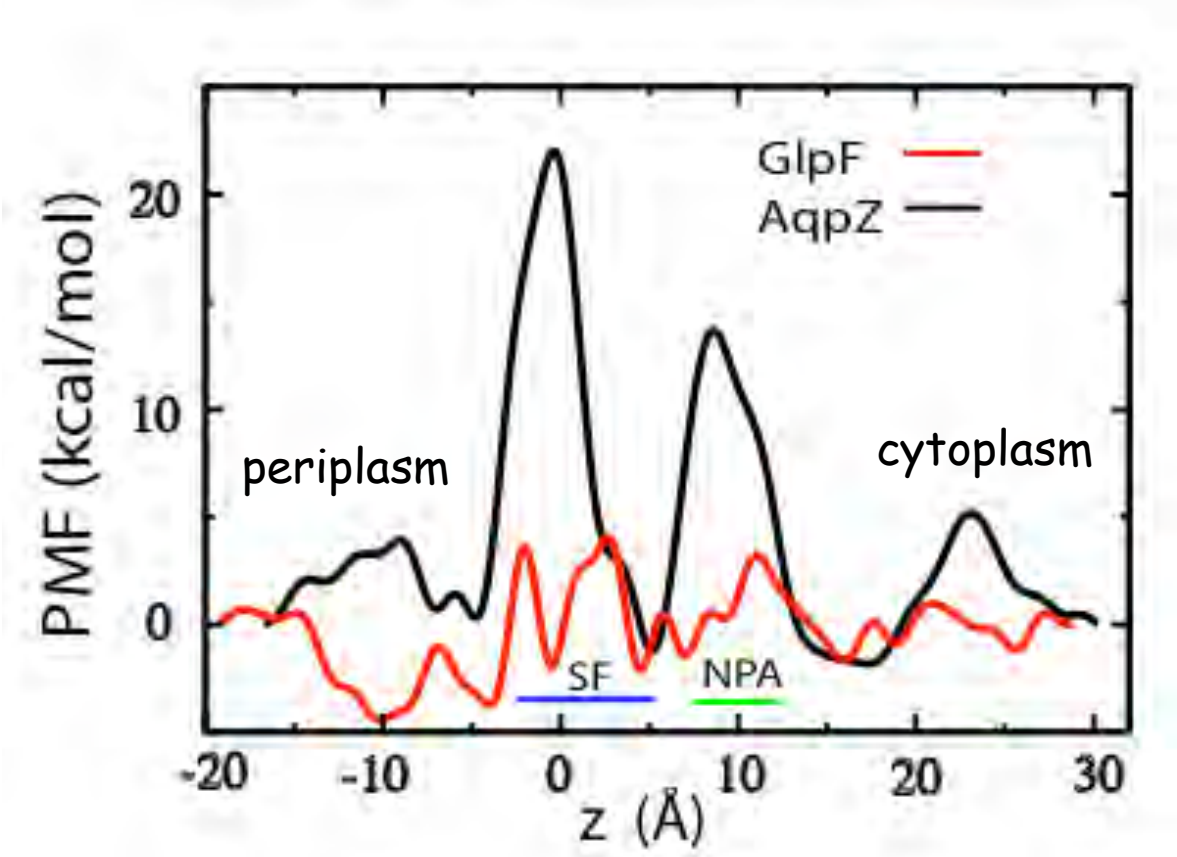
Jensen et al., *PNAS*, 99:6731-6736, 2002.

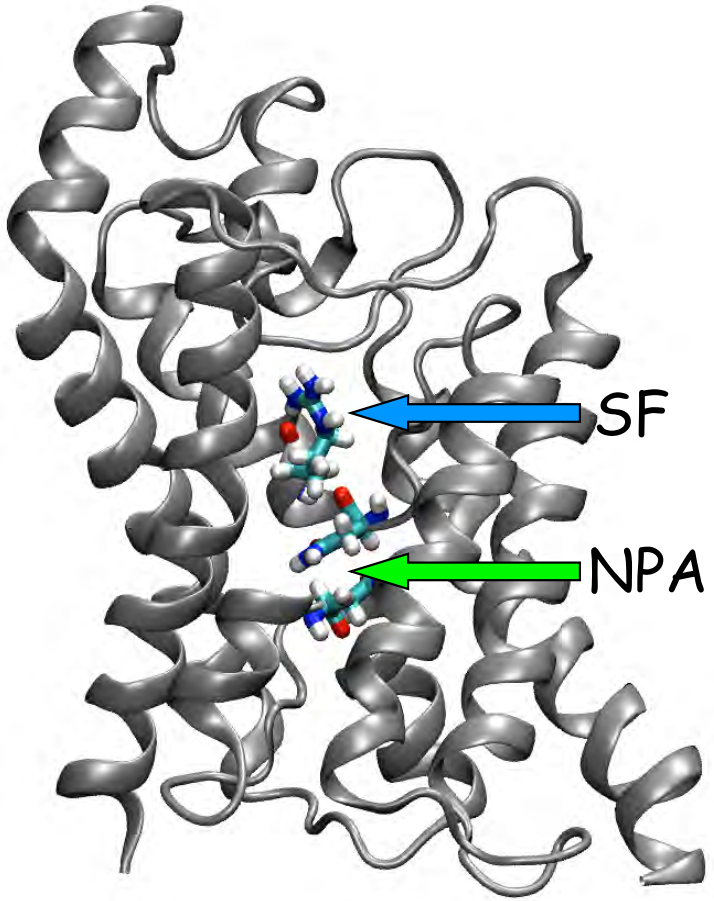
Artificial induction of glycerol conduction through AqpZ



Y. Wang, K. Schulten, and E. Tajkhorshid ***Structure*** 13, 1107 ***(***2005)

Three fold higher barriers



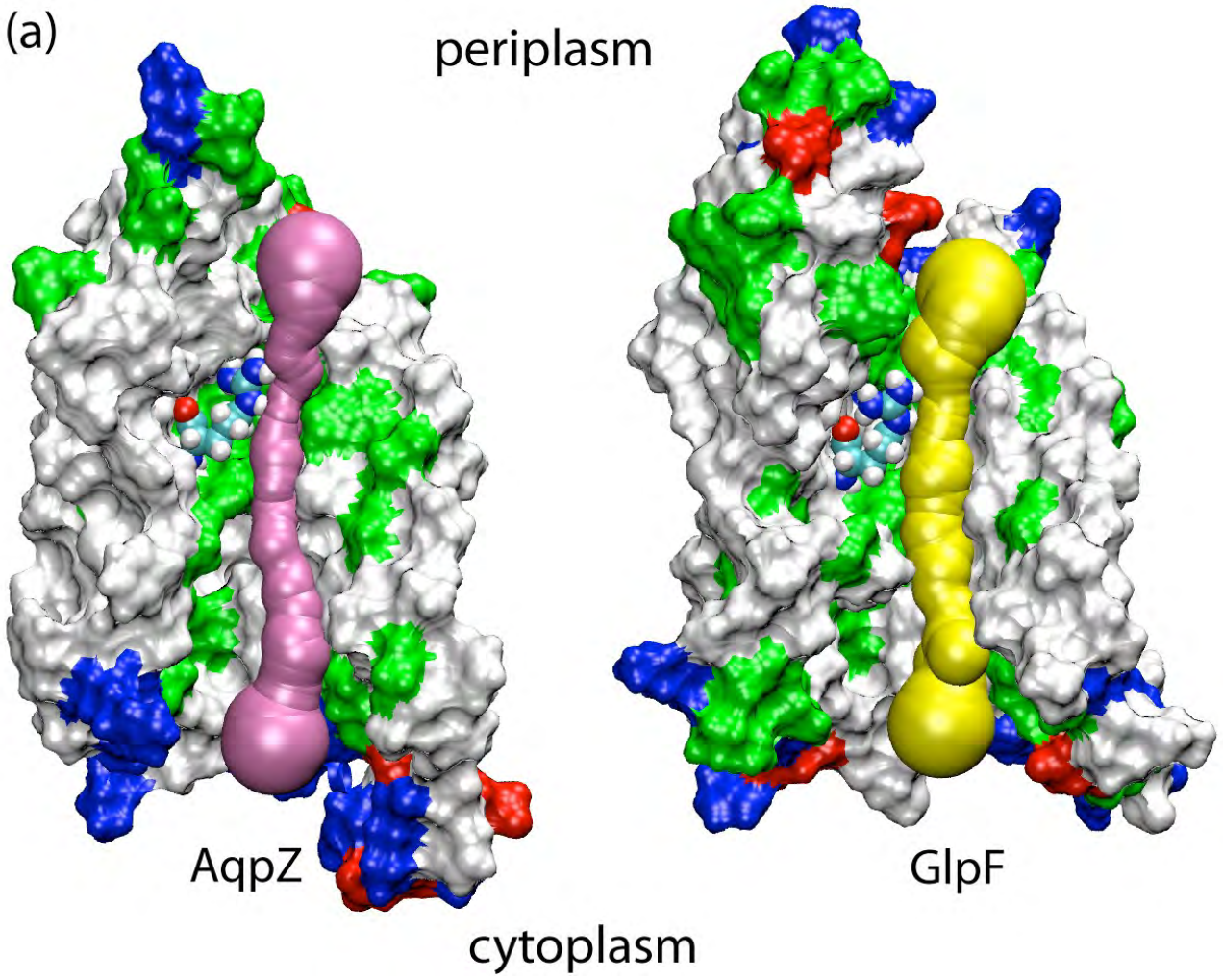


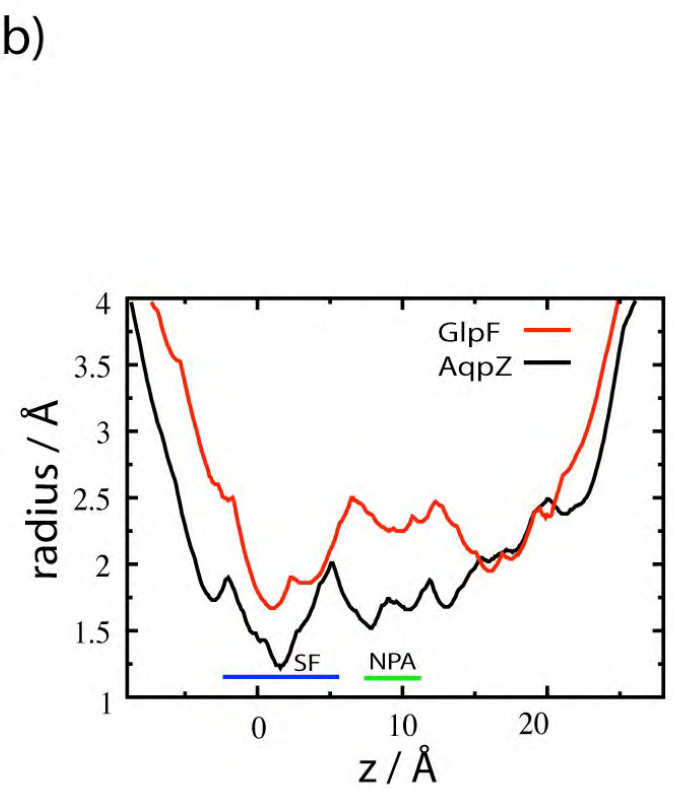
AqpZ 22.8 kcal/mol

GlpF 7.3 kcal/mol

Y. Wang, K. Schulten, and E. Tajkhorshid ***Structure*** 13, 1107 ***(***2005)

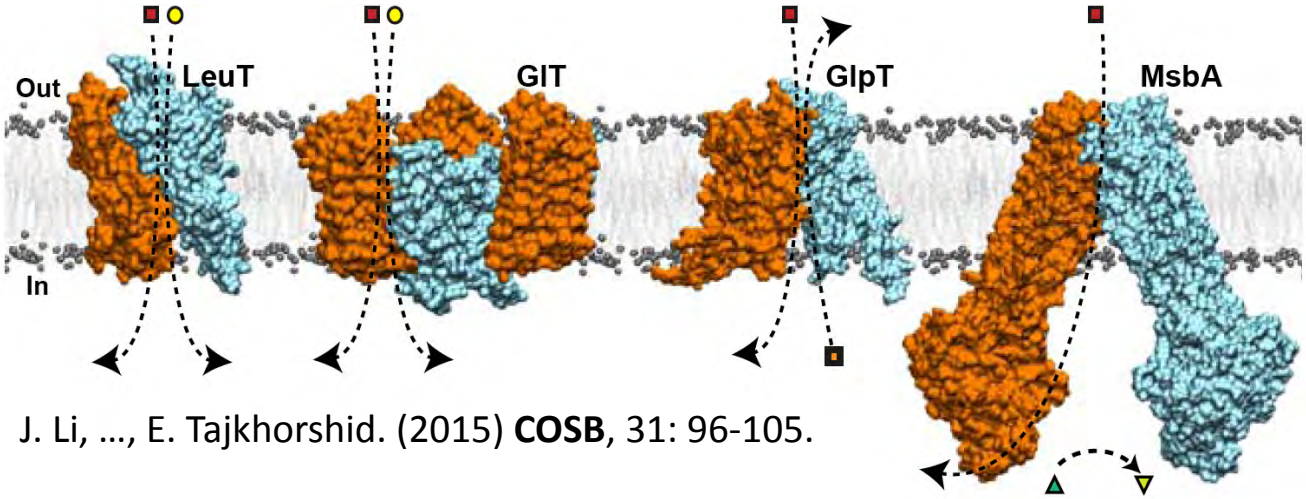
Could it be simply the size?



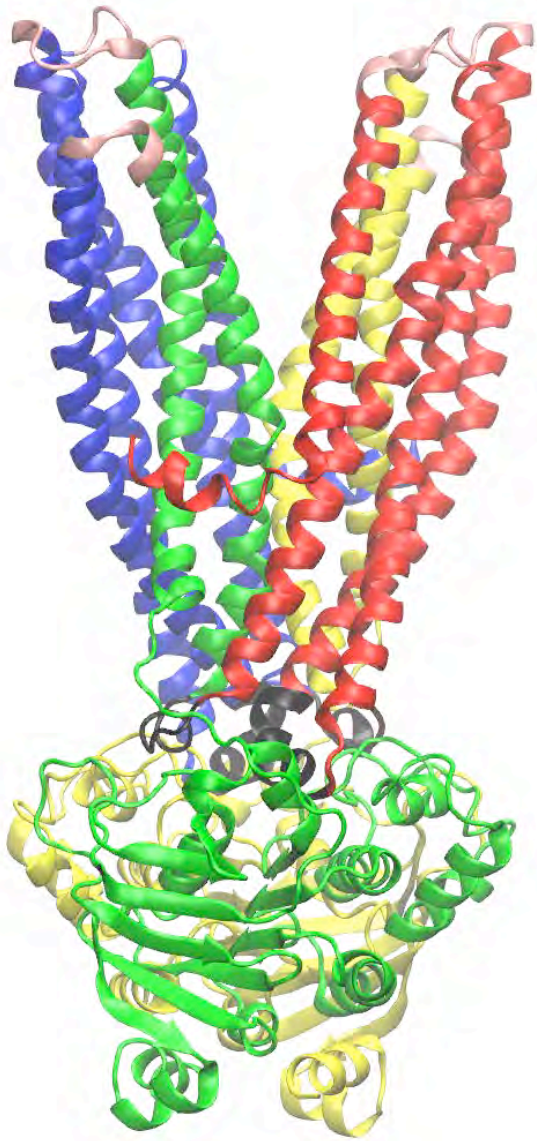


Y. Wang, K. Schulten, and E. Tajkhorshid ***Structure*** 13, 1107 ***(***2005)

Battling the Timescale - Case II
Biased (nonequilibrium) simulations



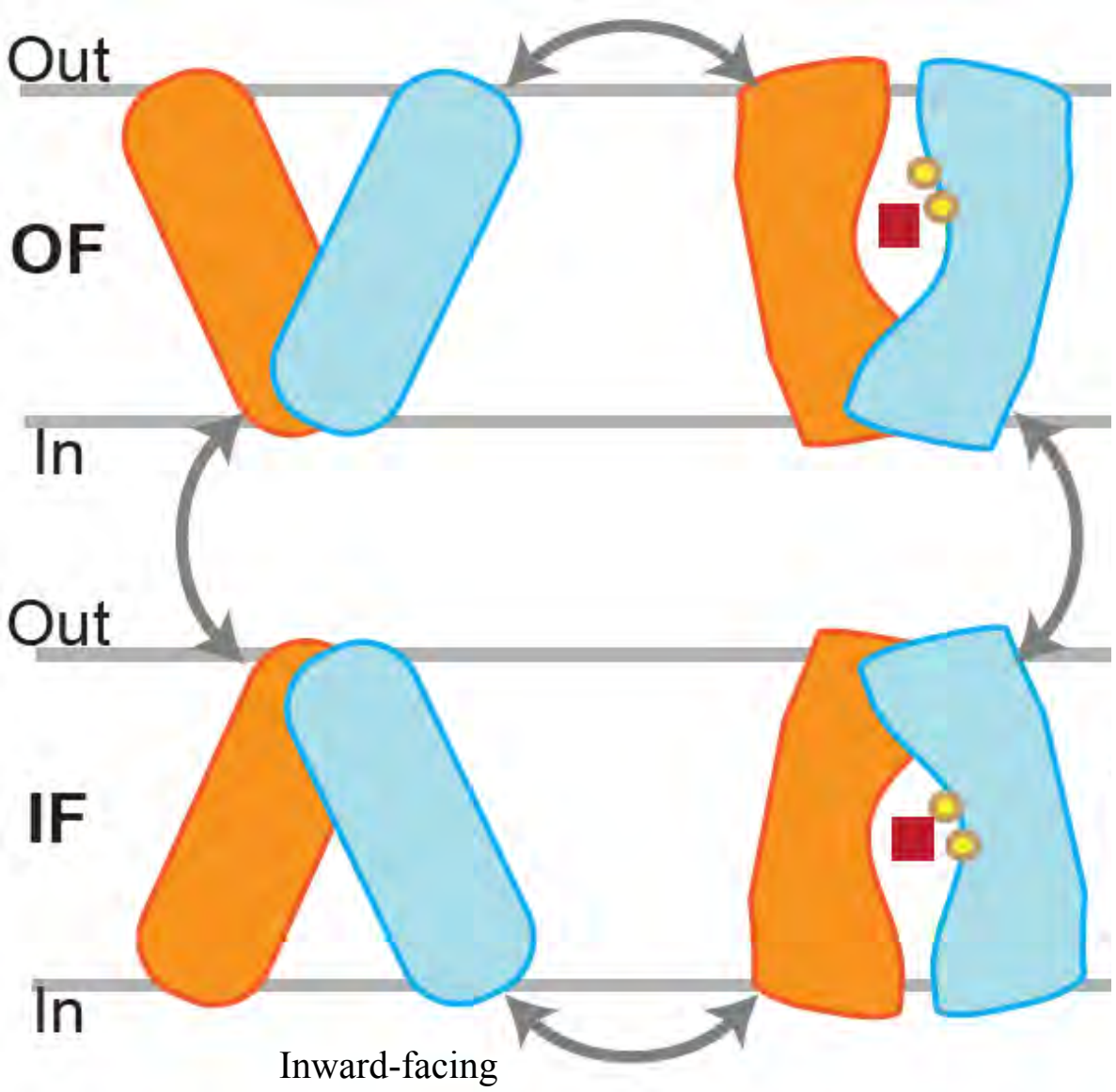
J. Li, …, E. Tajkhorshid. (2015) **COSB**, 31: 96-105.



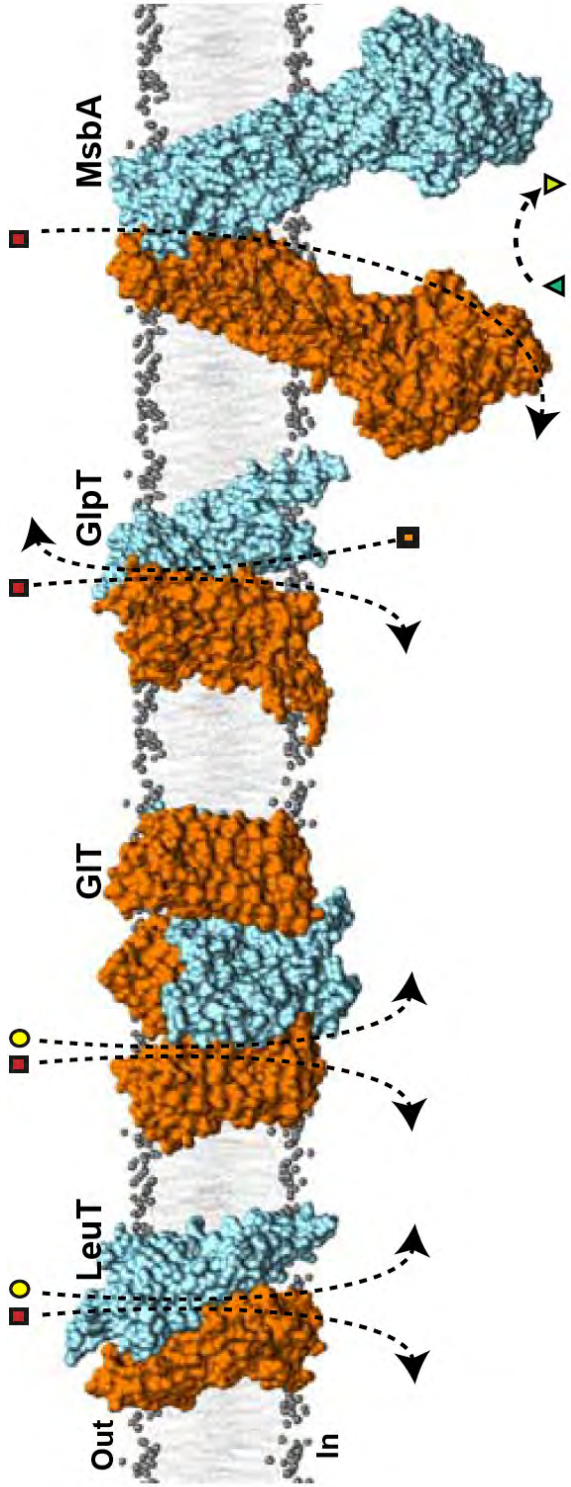
- **Neurotransmitter Uptake**
  - » Norepinephrine, serotonin, dopamine, glutamate,…
- **Gastrointestinal Tract**
  - » Active absorption of nutrients
  - » Secretion of ions
- **Kidneys**
  - » Reabsorption
  - » Secretion
- **Pharmacokinetics of all drugs**
  - » Absorption, distribution, elimination
  - » Multi-drug resistance in cancer cells

Alternating Access Mechanism

Outward-facing



Inward-facing

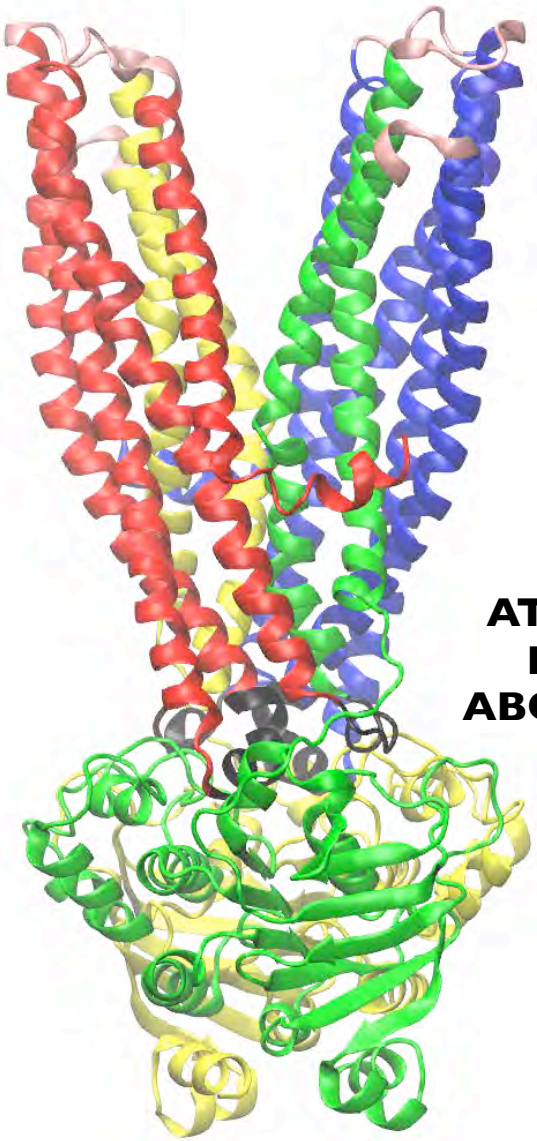


Jardetzky O. *Nature* **211**: 969–970 (1966)

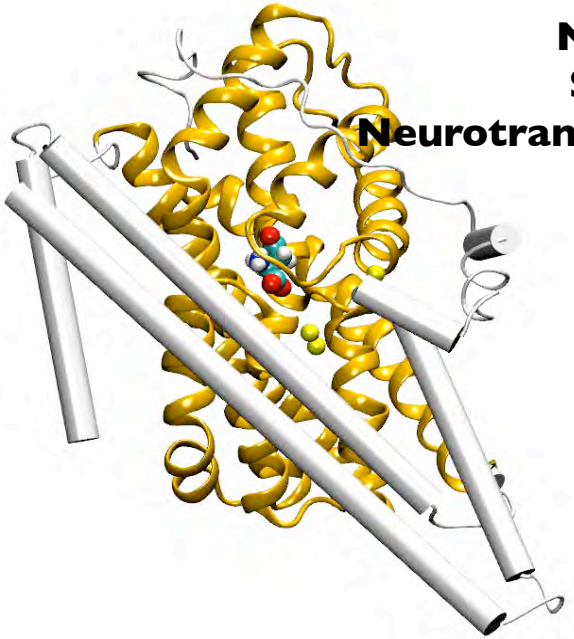
J. Li, …, E. Tajkhorshid. (2015) **COSB**, 31: 96-105.

COMPLEX

Diverse Structural Transitions Involved

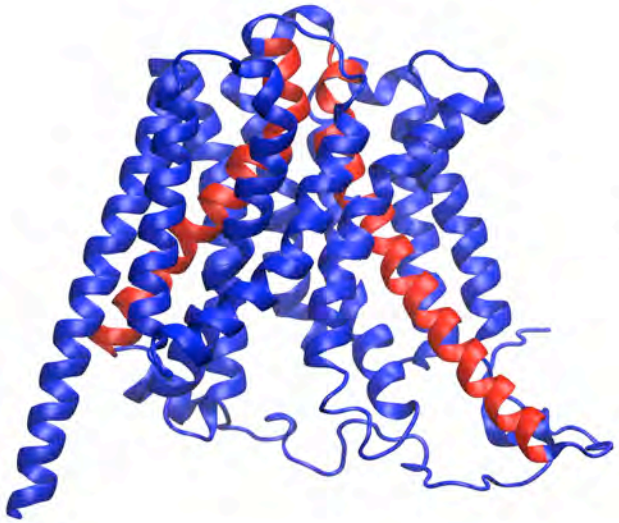


ATP-Driven
Primary
ABC Exporter



Na-coupled
Secondary
Neurotransmitter Transporter

Secondary
Phosphate
Antiporter



NON-EQUILIBRIUM METHODS ARE REQUIRED.

Complex Processes Require Complex Treatments

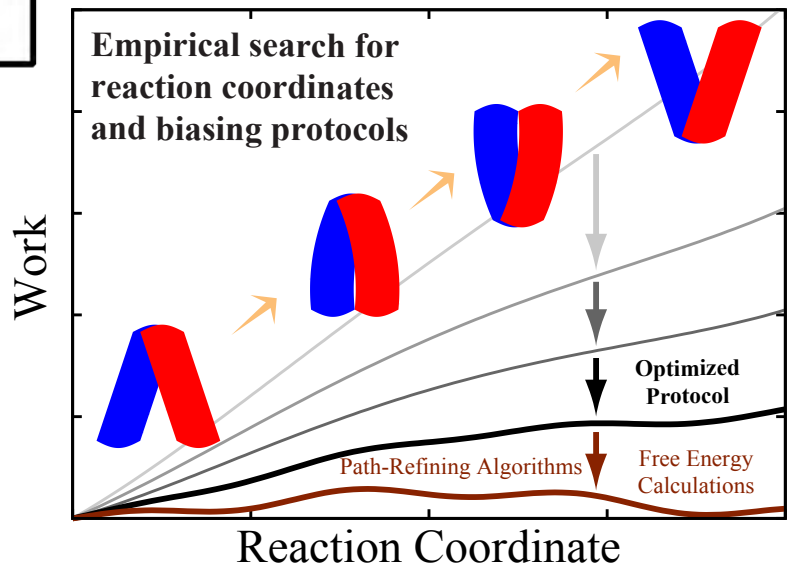
I.1 Defining Practical Collective Variables
Empirical search for practical collective variables for inducing the conformational changes involved in the transition.

I.2 Optimizing the Biasing Protocols
Systematic search for a practical biasing protocol by using different combinations of collective variables.

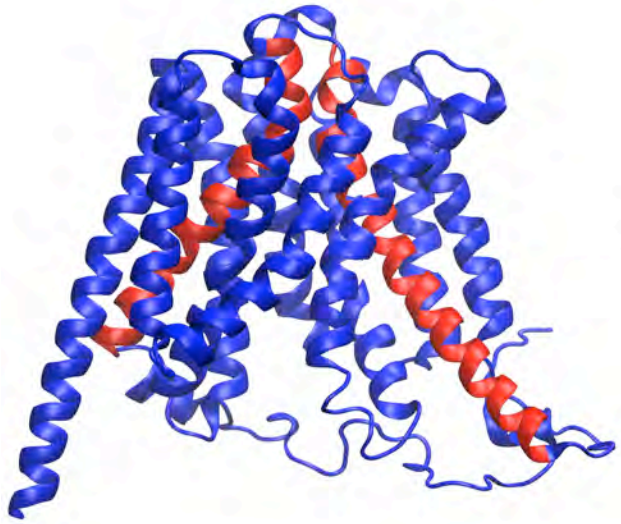
II. Optimizing the Transition Pathway
Use all of the conformations available to generate the most reliable transition pathway:
1. Bayesian approach for combining the data
2. Post-hoc string method (analysis tool)
3. String method with swarms of trajectories

III.1 Free Energy Calculations
Using the most relevant collective variables (from I.1), biasing protocol (from I.2), and initial conformations (from I.2).

III.2 Assessing the Sampling Efficiency
Detecting the poorly sampled, but potentially important regions, e.g., by using PCA.

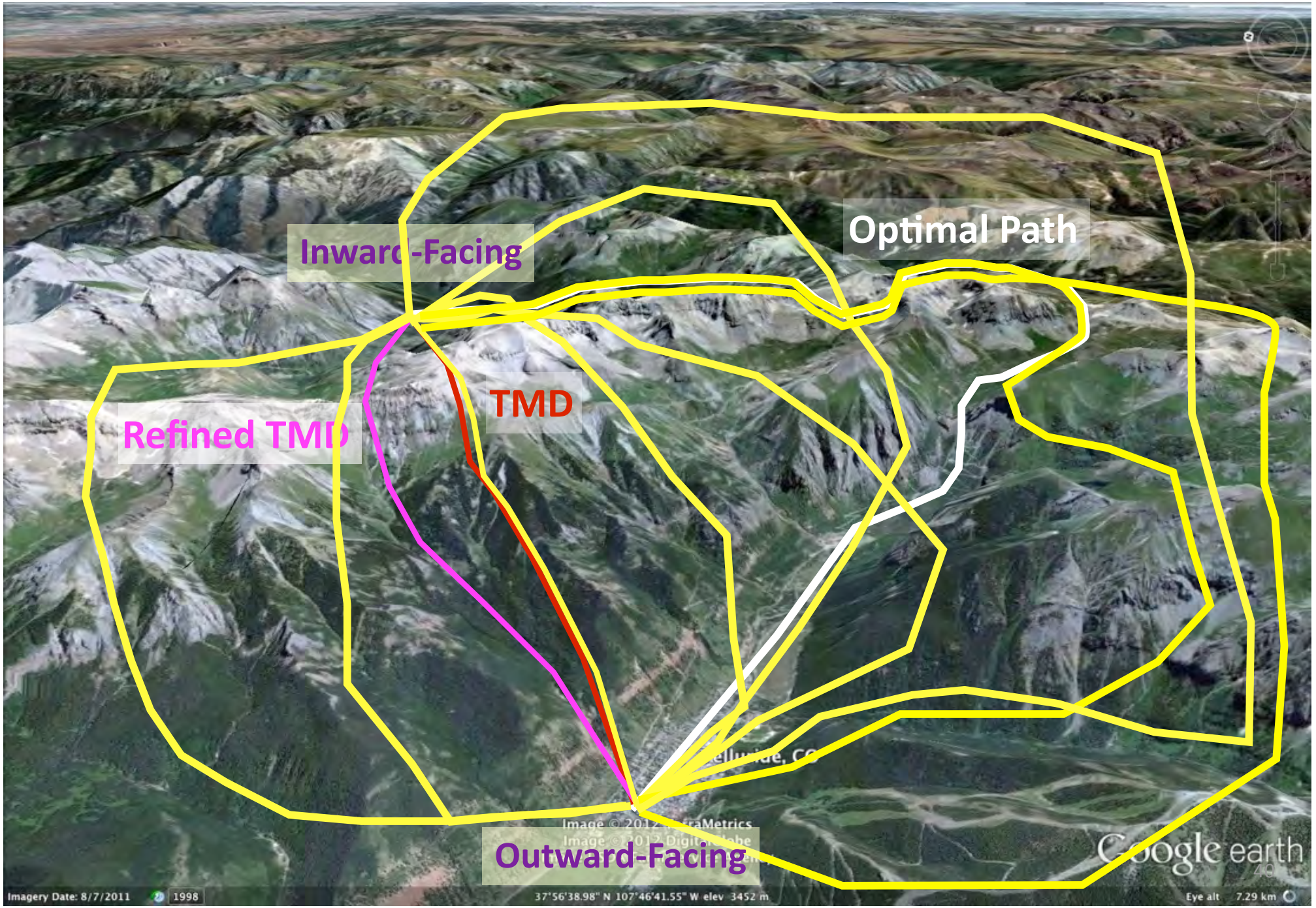


Mahmoud Moradi



M. Moradi and ET (2013) **PNAS**, 110:18916–18921.
M. Moradi and ET (2014) **JCTC**, 10: 2866–2880.
M. Moradi, G. Enkavi, and ET (2015) **Nature Comm.**, 6:8393.

Aggressive Search of the Space



Non-equilibrium Driven Molecular Dynamics:

Applying a time-dependent external force to induce the transition

Along various pathways/mechanisms (collective variables)

Harmonic constant

Initial state

$$U_{dr}(\boldsymbol{x}, t) = \frac{1}{2}k\left(\xi(\boldsymbol{x}) - \xi_A + (\xi_B - \xi_A)\frac{t}{\mathrm{T}}\right)^2$$

Biasing potential

Collective variables:
RMSD, distance,
R_g, angle, …
orientation quaternion

Final state

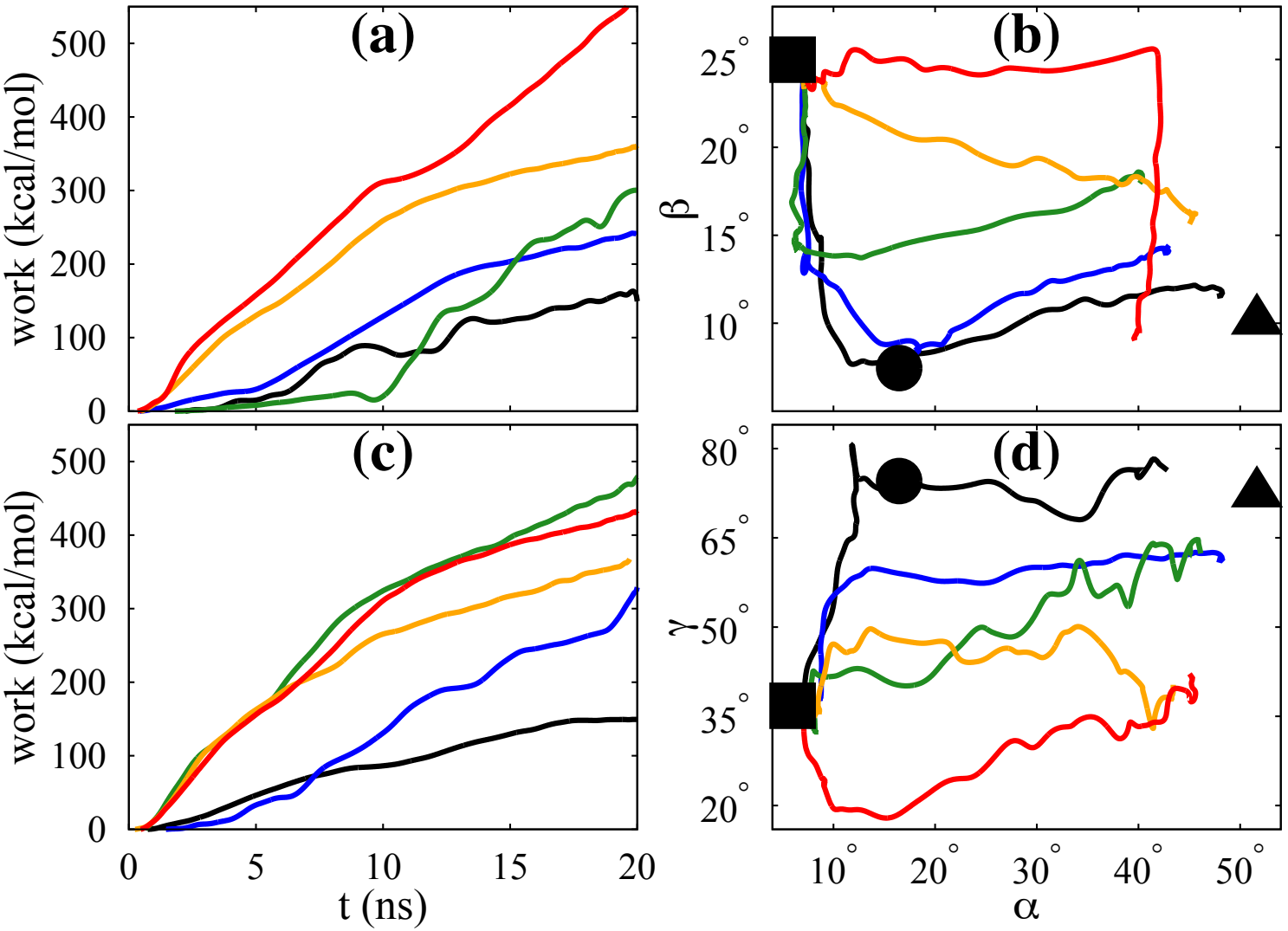
Total simulation time

M. Moradi and ET (2013) **PNAS**, 110:18916–18921.
M. Moradi and ET (2014) **JCTC**, 10: 2866–2880.
M. Moradi, G. Enkavi, and ET (2015) **Nature Comm.**, 6:8393.

Progressively Optimizing the Biasing Protocol/Collective Variable using non-Equilibrium Work as a Measure of the Path Quality

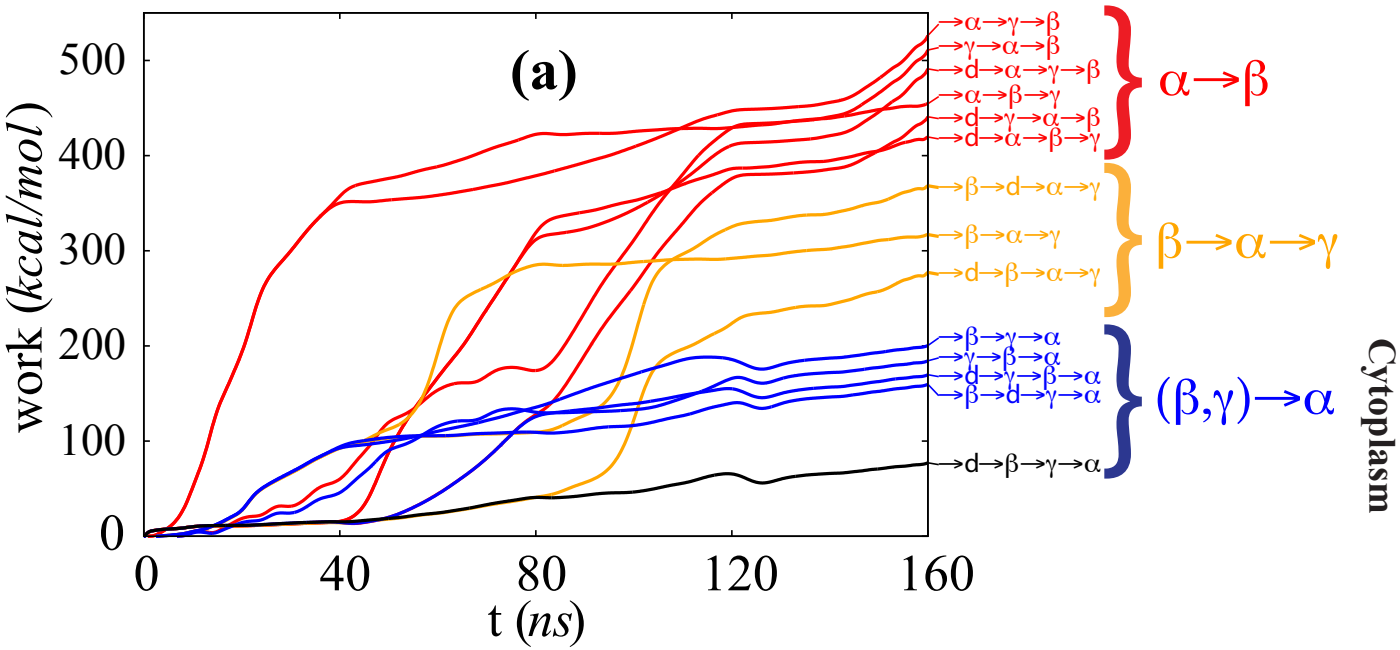
Work

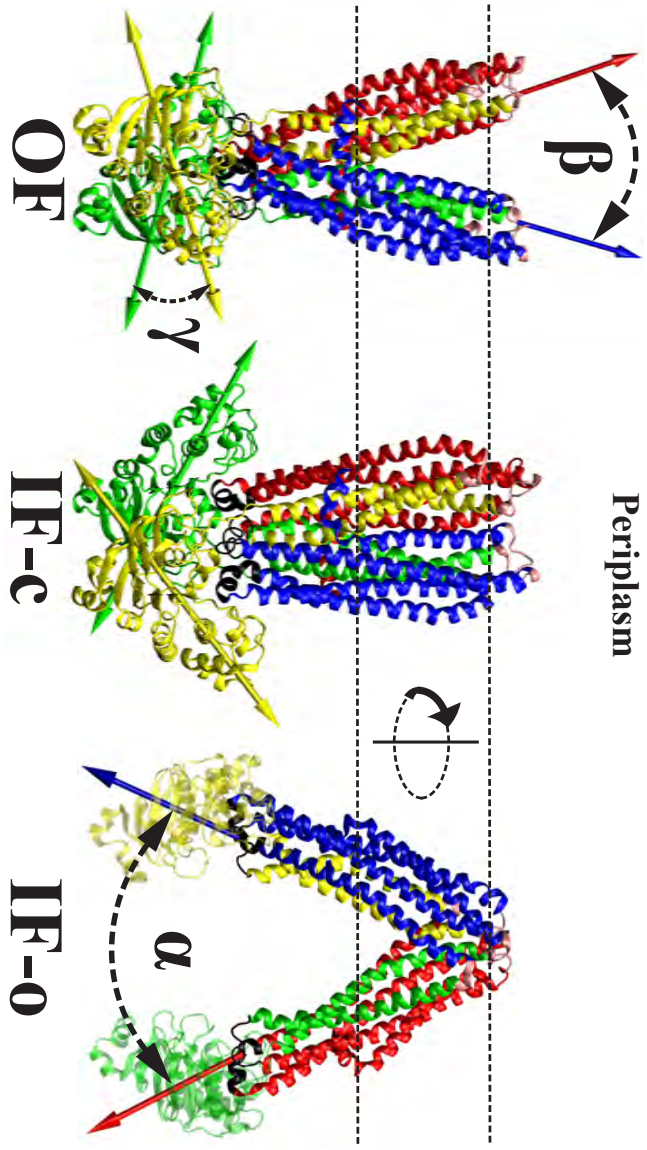
Mechanism



Example set taken from a subset of 20 ns biased simulations

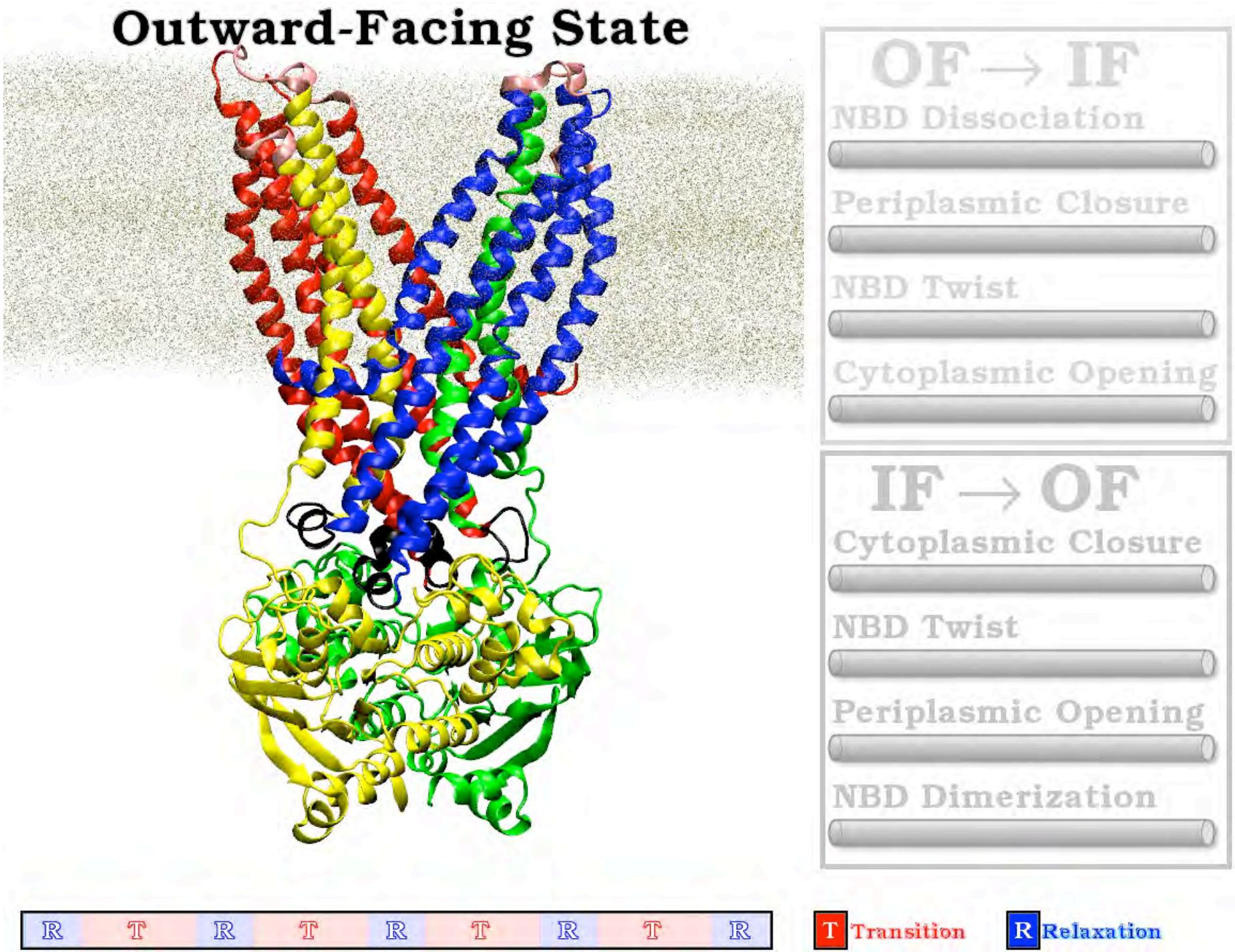
Mechanistic Insight From Transition Pathways in ABC exporters from Non-Equilibrium Simulations





M. Moradi and ET (2013) **PNAS**, 110:18916–18921.

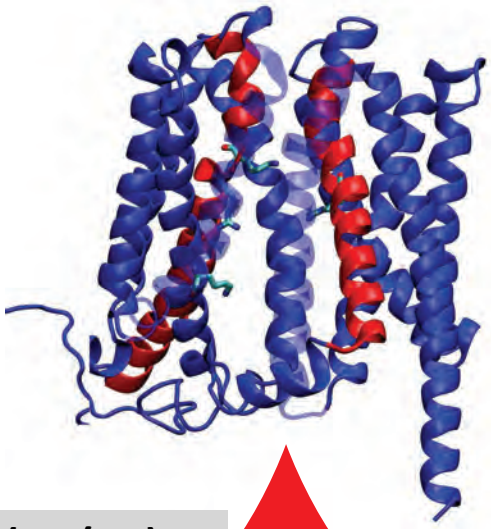
M. Moradi and ET (2014) **JCTC**, 10: 2866–2880.



NBD Doorknob Mechanism

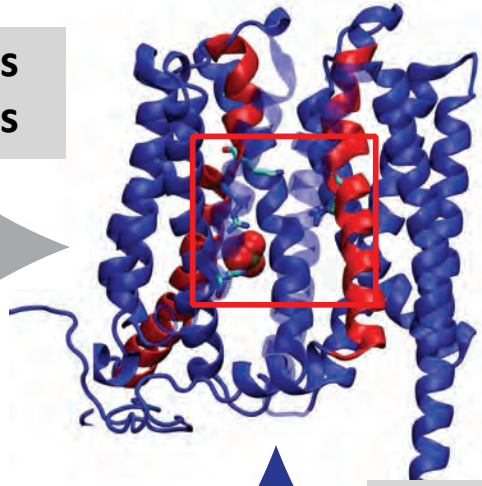
M. Moradi and ET (2013) **PNAS**, 110:18916–18921.

Describing a Complete Cycle (Adding Substrate) Requiring a Combination of Multiple Collective Variables



IF_{apo}

30 r x 20 ns
30 r x 20 ns

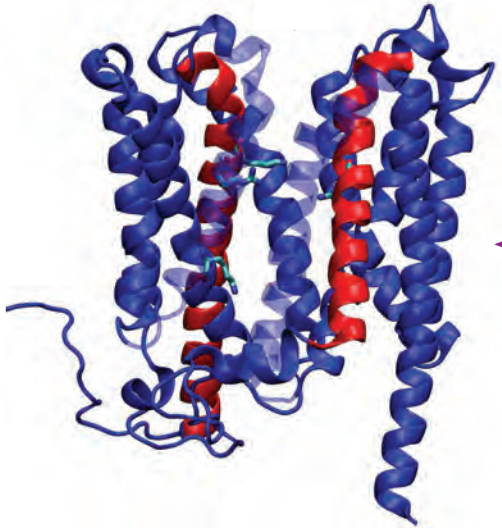


IF_{bound}

12 replicas x 40 ns (H1/H7)
50 replicas x 20 ns (10 Hs)

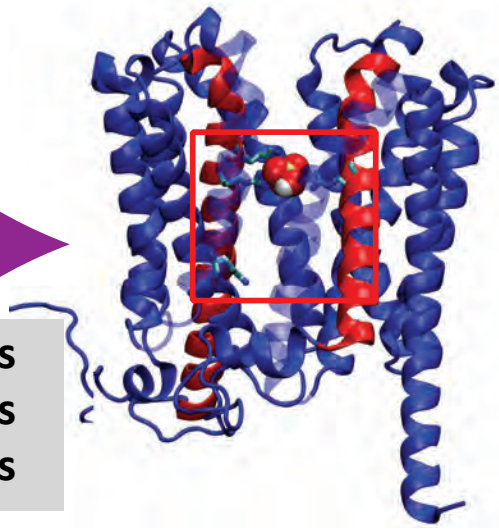
150 replicas

12 replicas x 40 ns (H1/H7)
24 replicas x 20 ns (H1/H7)
200 replicas (2D) x 5 ns
50 replicas x 20 ns



OF_{apo}

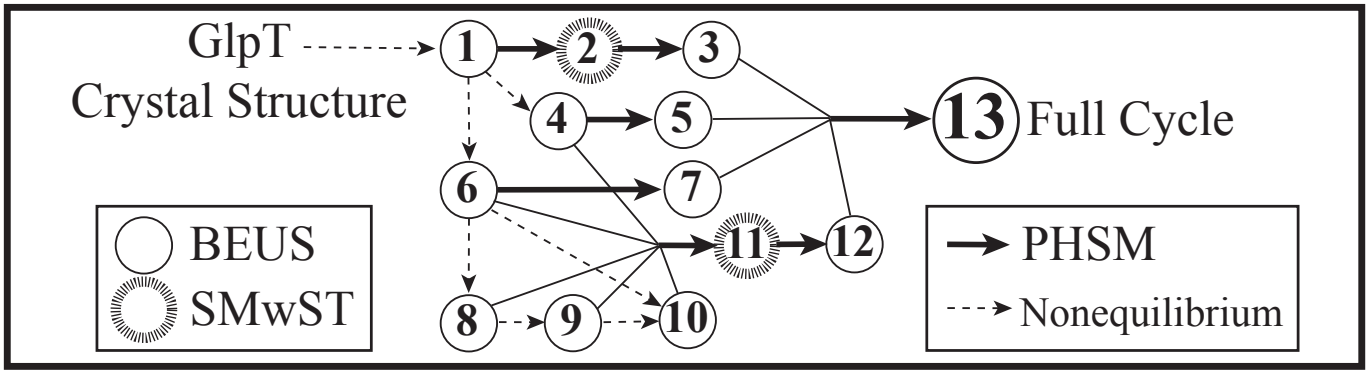
30 r x 20 ns
30 r x 20 ns
30 r x 20 ns



OF_{bound}

Simulation protocols

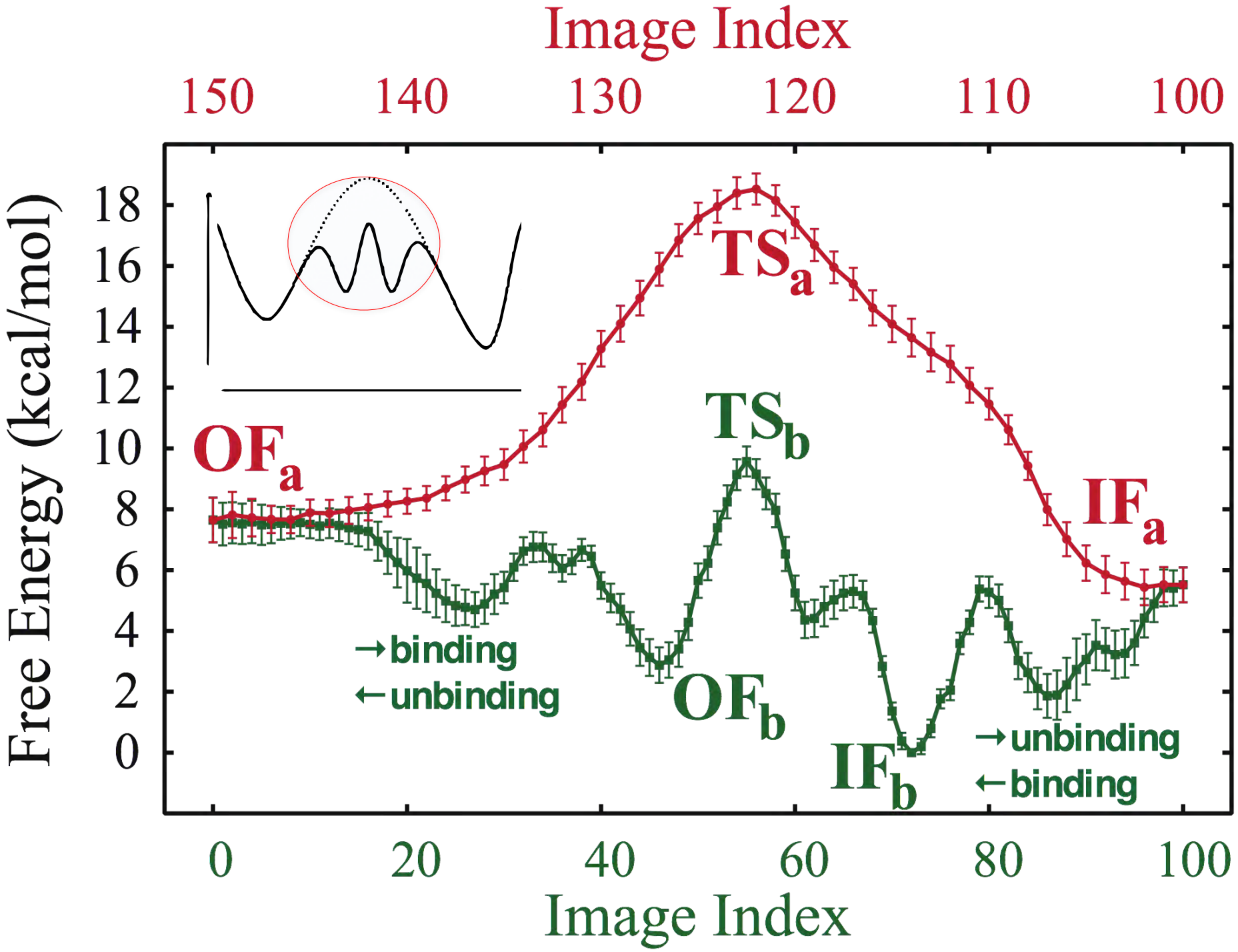
	Transition	Technique	Collective Variables	# of Replicas × Runtime		
1	$IF_a \leftrightarrow OF_a$	BEUS	(Q_1,Q_7)	12 × 40 ns	=	0.5 µs
2		SMwST	{Q}	1000 × 1 ns	=	1 µs
3		BEUS	{Q}	50 × 20 ns	=	1 µs
4	$IF_a \leftrightarrow IF_b$	BEUS	Z_{Pi}	30 × 40 ns	=	1.2 µs
5		BEUS	$(\{Q\}, Z_{Pi})$	30 × 40 ns	=	1.2 µs
6	$OF_a \leftrightarrow OF_b$	BEUS	Z_{Pi}	30 × 40 ns	=	1.2 µs
7		BEUS	$(\{Q\}, Z_{Pi})$	30 × 40 ns	=	1.2 µs
8	$IF_b \leftrightarrow OF_b$	BEUS	(Q_1,Q_7)	24 × 20 ns	=	0.5 µs
9		BEUS	Z_{Pi}	15 × 30 ns	=	0.5 µs
10		2D BEUS	$(\Delta RMSD, Z_{Pi})$	200 × 5 ns	=	1 µs
11		SMwST	$(\{Q\}, Z_{Pi})$	1000 × 1 ns	=	1 µs
12		BEUS	$(\{Q\}, Z_{Pi})$	50 × 20 ns	=	1 µs
13	**Full Cycle**	**BEUS**	**$(\{Q\}, Z_{Pi})$**	**150 × 50 ns**	=	**7.5 µs**
	Total Simulation Time					**18.7 µs**



U.S. DEPARTMENT OF ENERGY
INCITE
LEADERSHIP COMPUTING

BLUE WATER NCSA

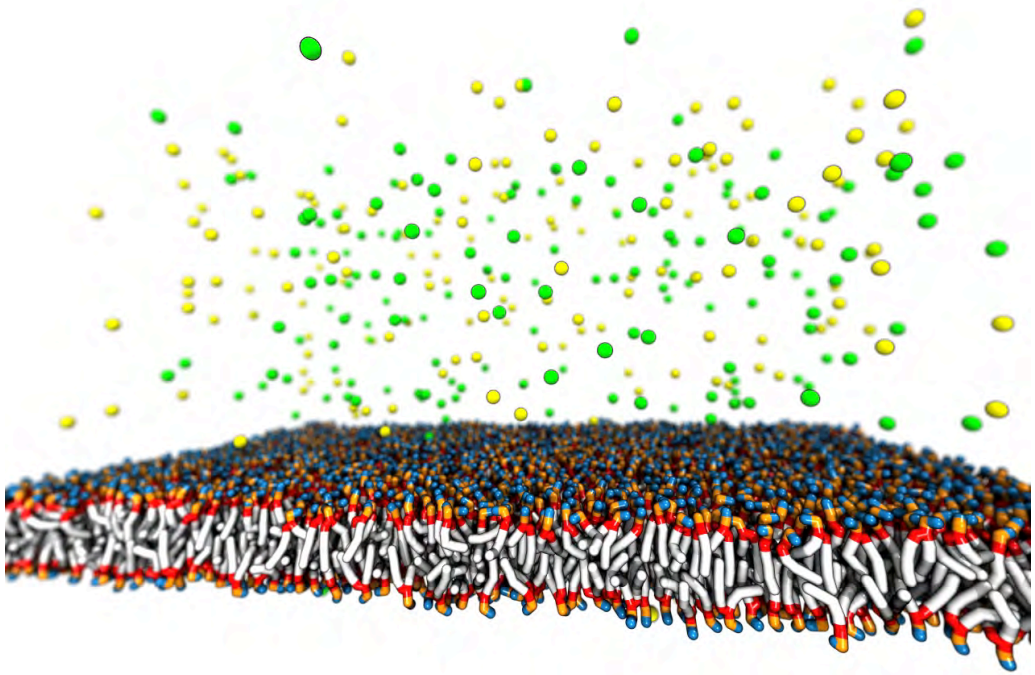
M. Moradi, G. Enkavi, and ET (2015) **Nature Communication**, 6: 8393.



M. Moradi, G. Enkavi, and ET (2015) **Nature Communication**, 6: 8393.

Battling the Timescale - Case III Multiscale Simulations

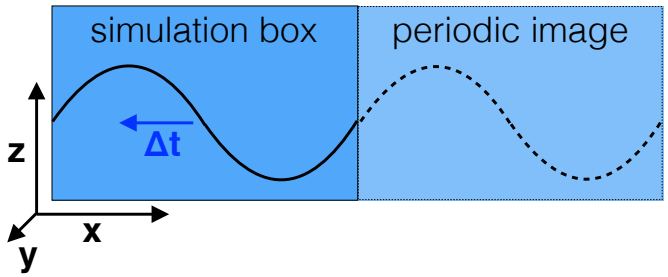
Membrane Budding/Fusion

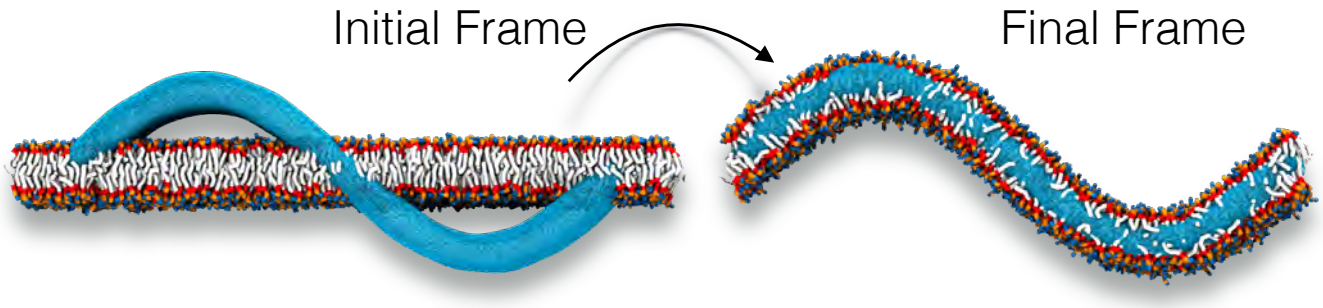


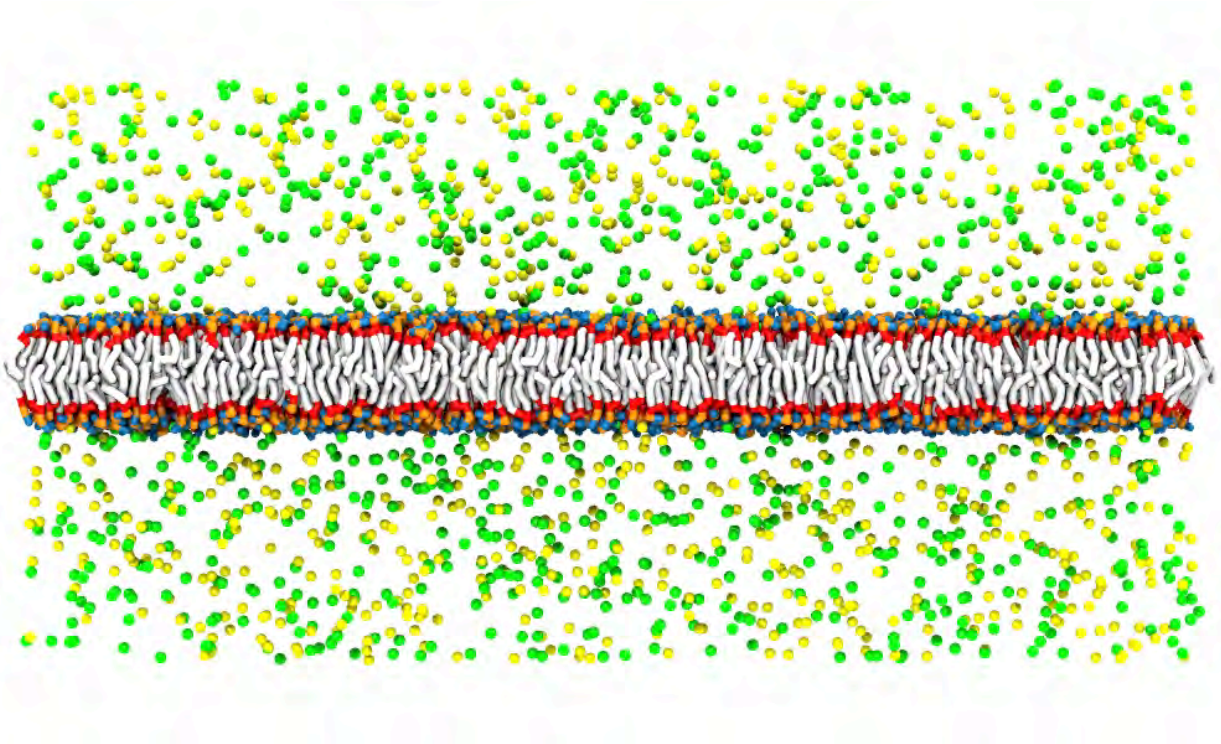
Combining multiple replica simulations and coarse-grained models to describe membrane fusion

Workflow for Multi-Scale Modeling

Parametrically Defined Sine Function

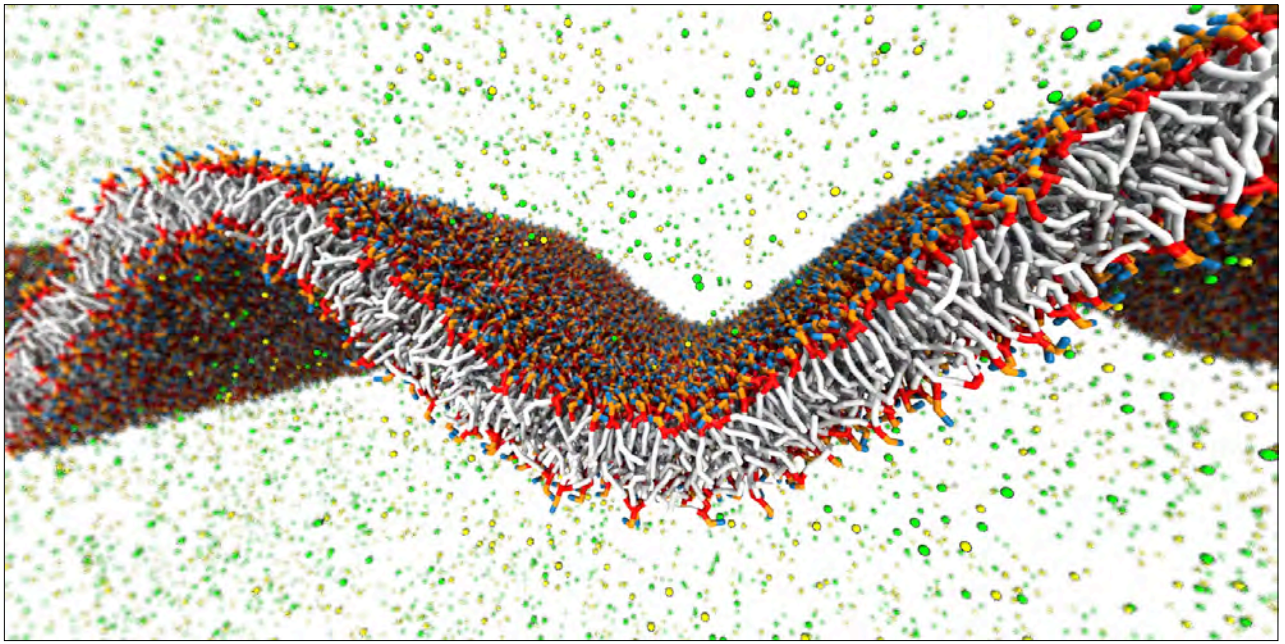






Christopher Mayne, Tajkhorshid Lab

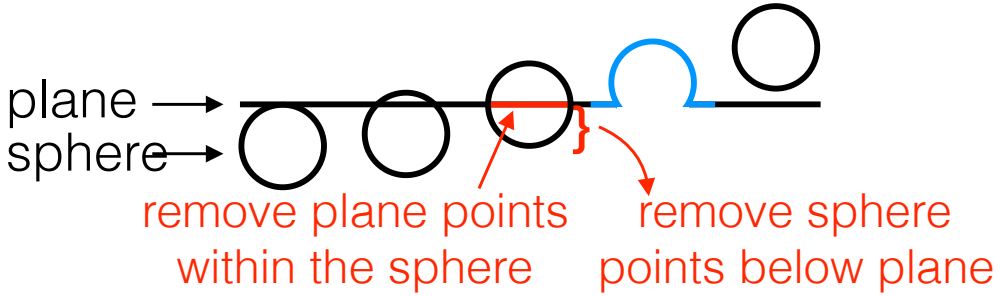
Workflow for Multi-Scale Modeling

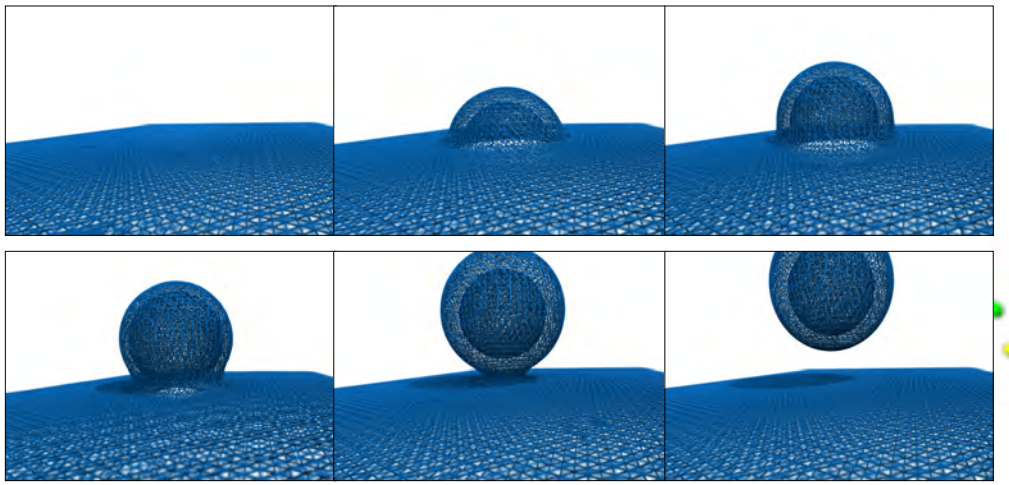


Christopher Mayne, Tajkhorshid Lab

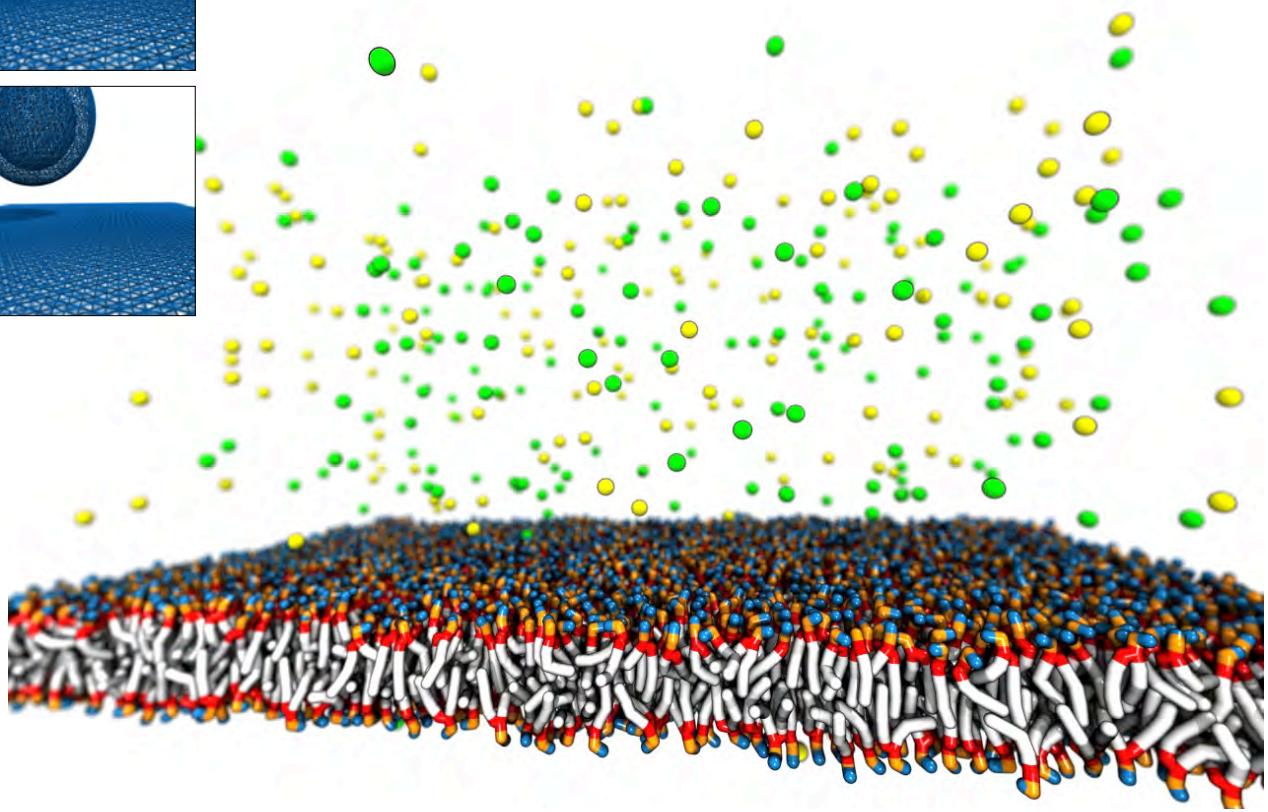
Workflow for Multi-Scale Modeling

Grid Design and Construction





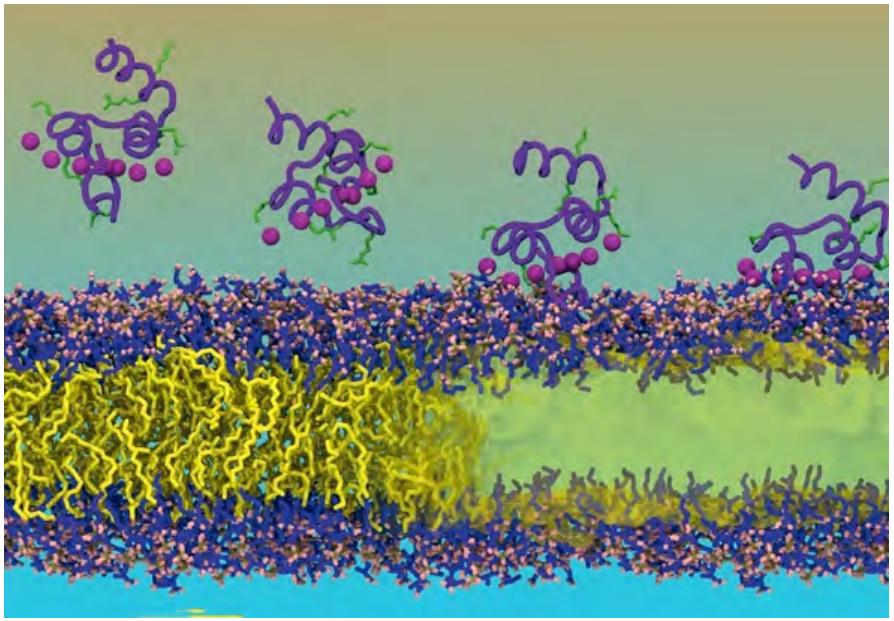
Membrane Budding/Fusion



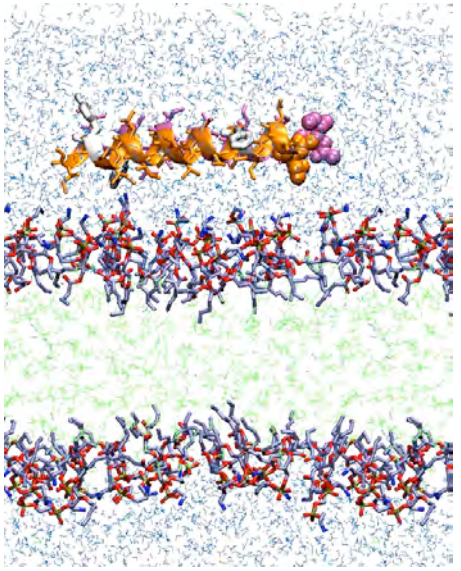
Christopher Mayne, Tajkhorshid Lab

Battling the Timescale - Case IV
Reduced Representations

Highly Mobile Membrane Mimetic model



GpA insertion in 12 ns



Specific lipids regulate various functional aspects of membrane proteins

Integral membrane proteins

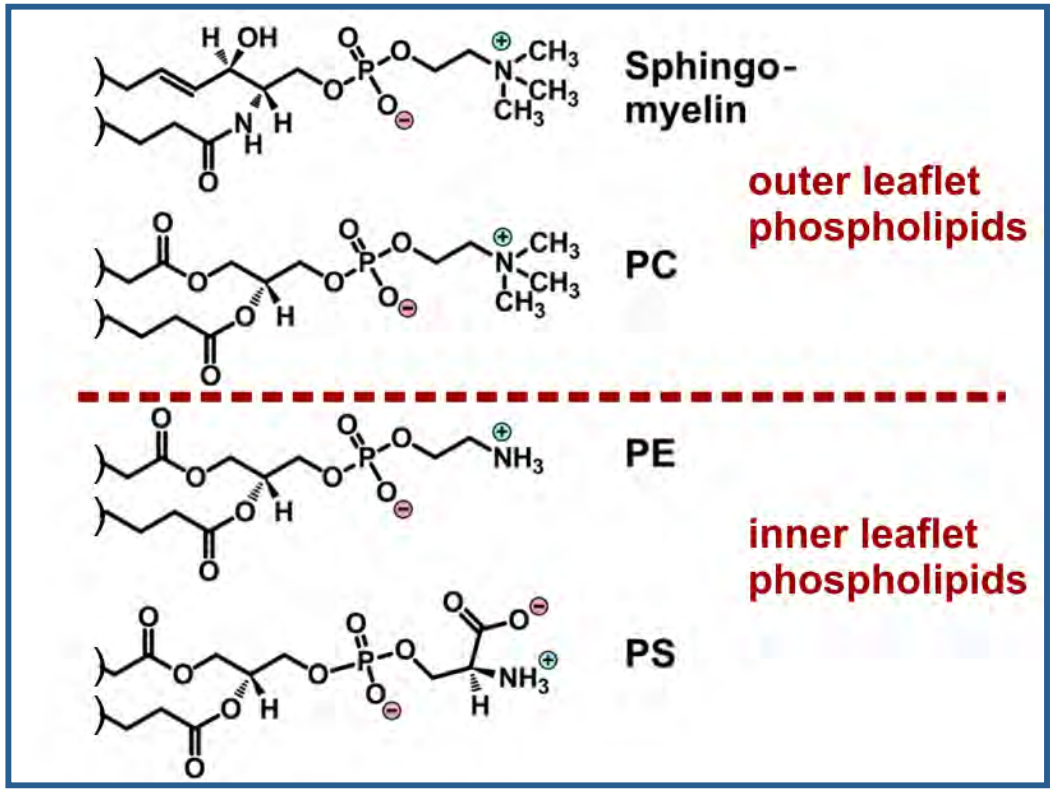
Peripheral membrane proteins

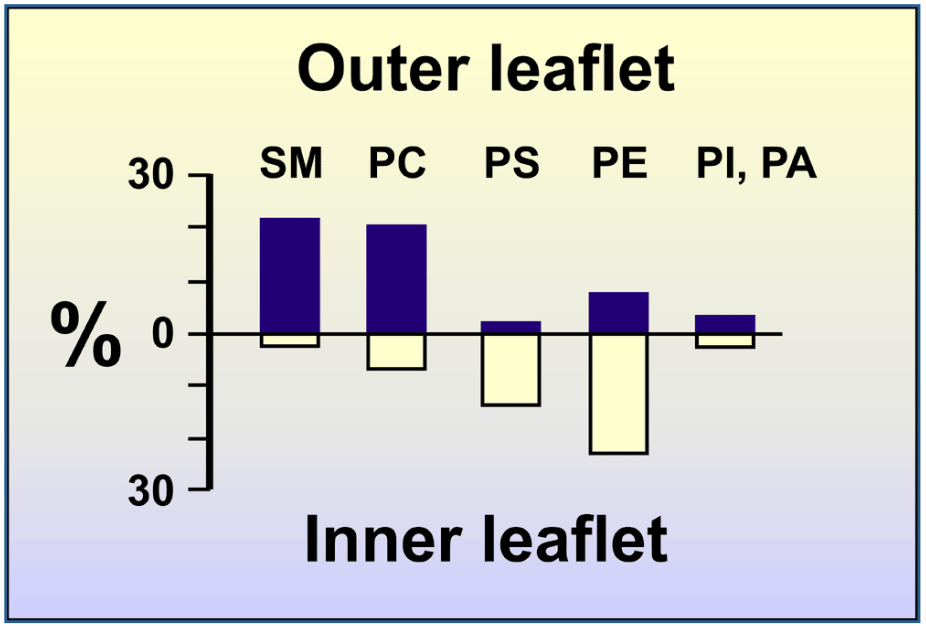
Lipid-Dependent Regulation and Activity of Peripheral Membrane Proteins

- Membrane binding is a key regulatory step in the function of diverse proteins:
  - Cytoplasmic enzymes (kinases, Ras, P450, synaptotagmin, …)
  - Coagulation factors (GLA and C2 domains)
  - Membrane sculpting proteins (BAR domain)
  - Pathogenic systems – viral fusion peptides, synuclein,
  - Immune/apoptotic system (TIM proteins)

- **Lipid-specificity** is a common feature:
  - Mostly at the level of **head groups**: PS, PG, PIP2, PA, …
  - Requiring all-atom representation of the head groups
  - Slow lateral diffusion of lipids within a bilayer environment makes simulation studies of membrane-associated phenomena even more challenging

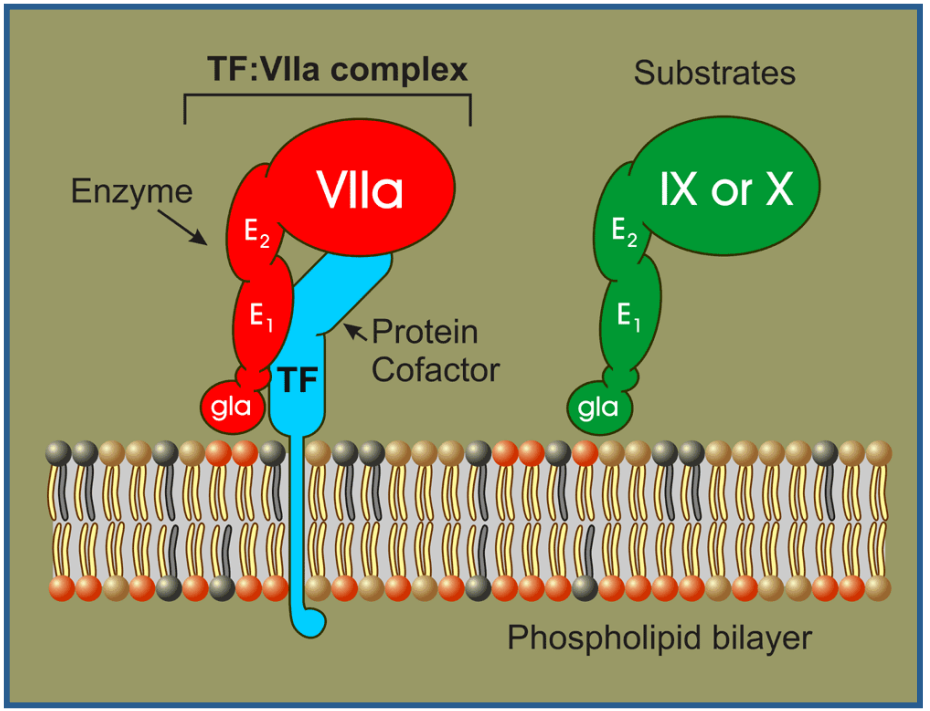
Lipid Dependent Binding and Activation





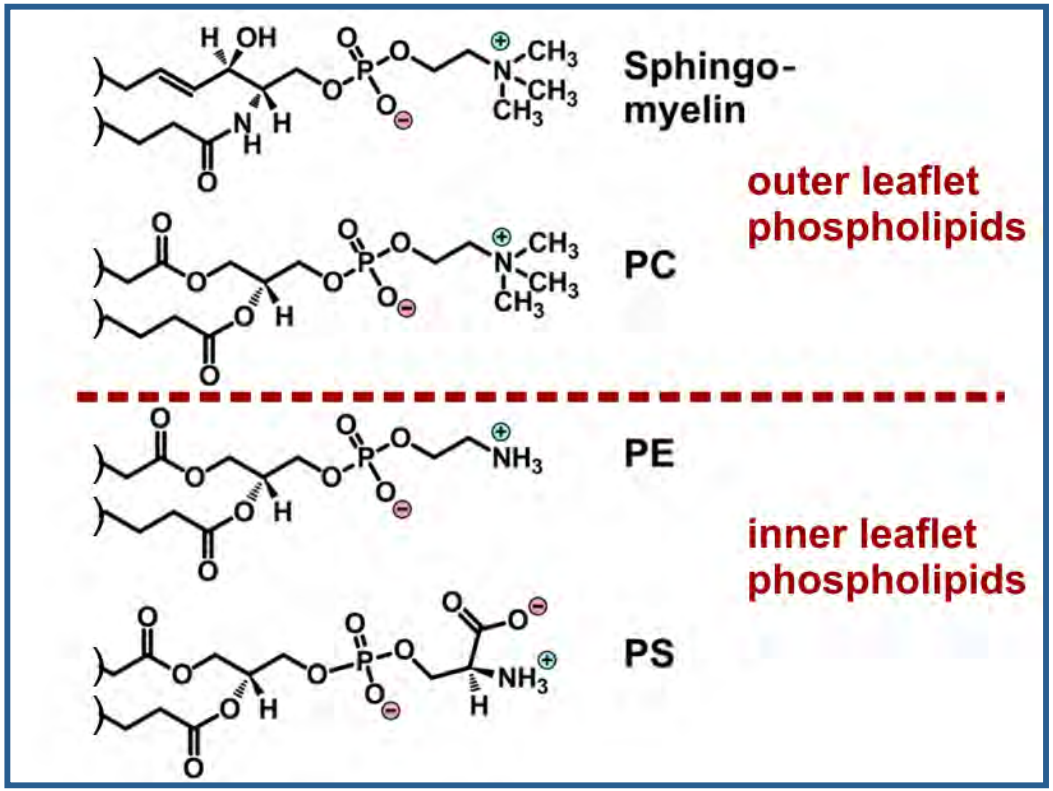
Affinity is controlled by lipid content

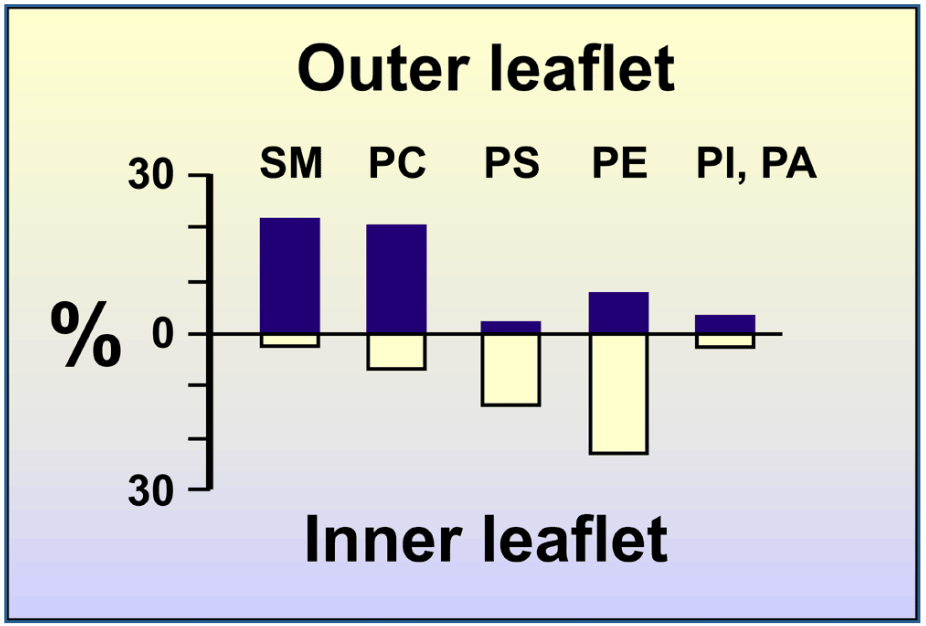
Leaflet asymmetry is vital for coagulation



Courtesy of Jim Morrissey, UIUC

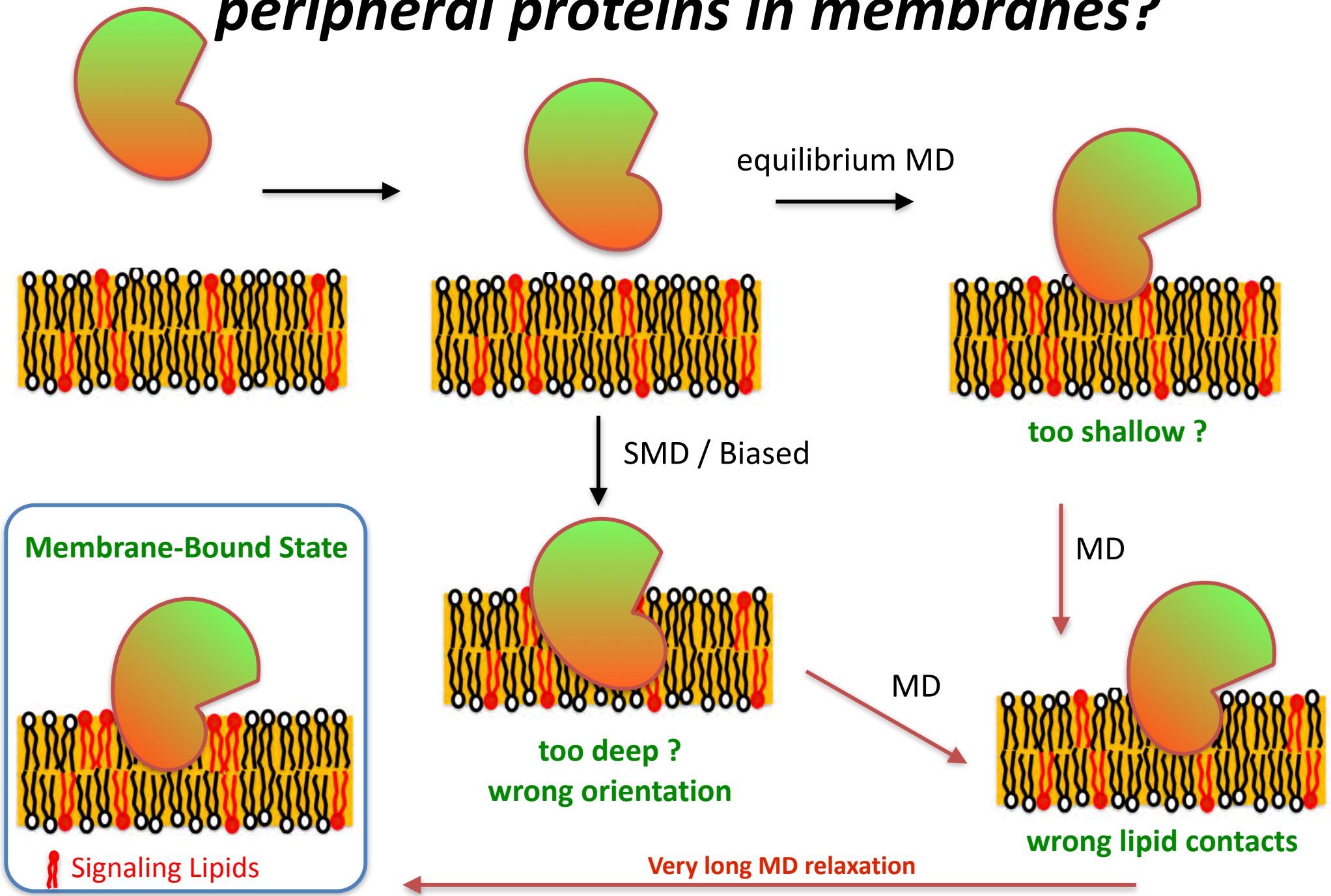
Lipid Dependent Binding and Activation



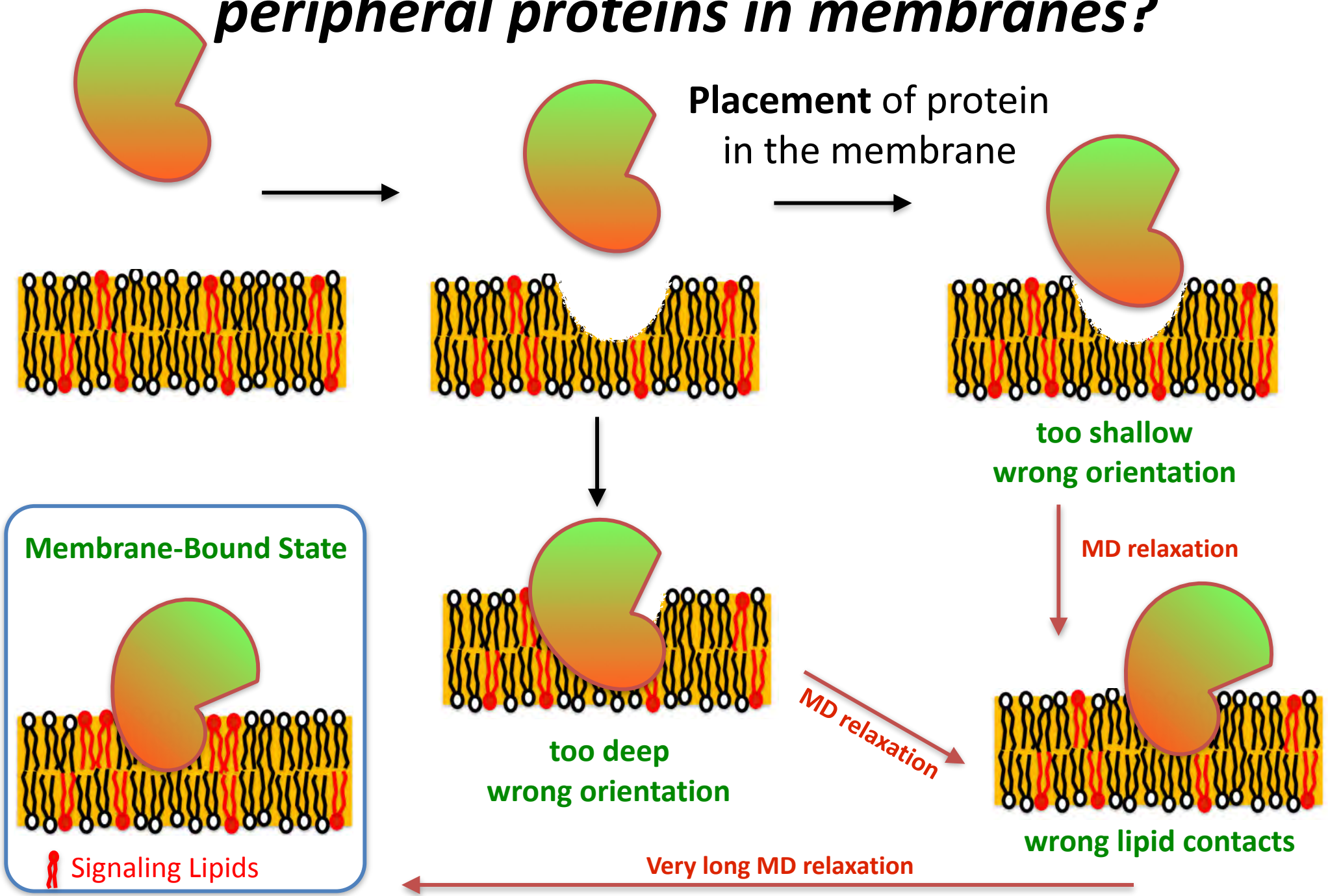


Mode and specificity of lipid-protein interactions constitute one of the main mechanistic aspects

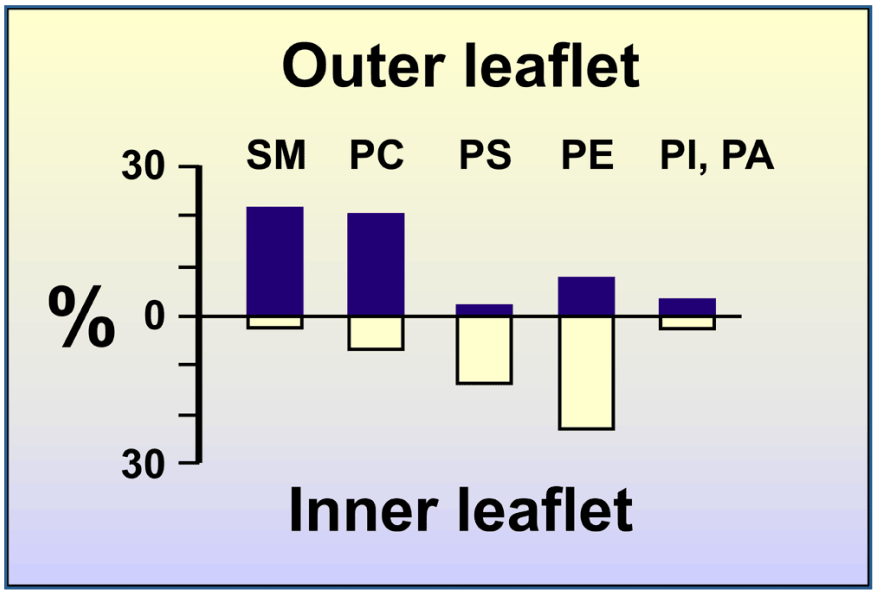
How do we construct an initial model for peripheral proteins in membranes?

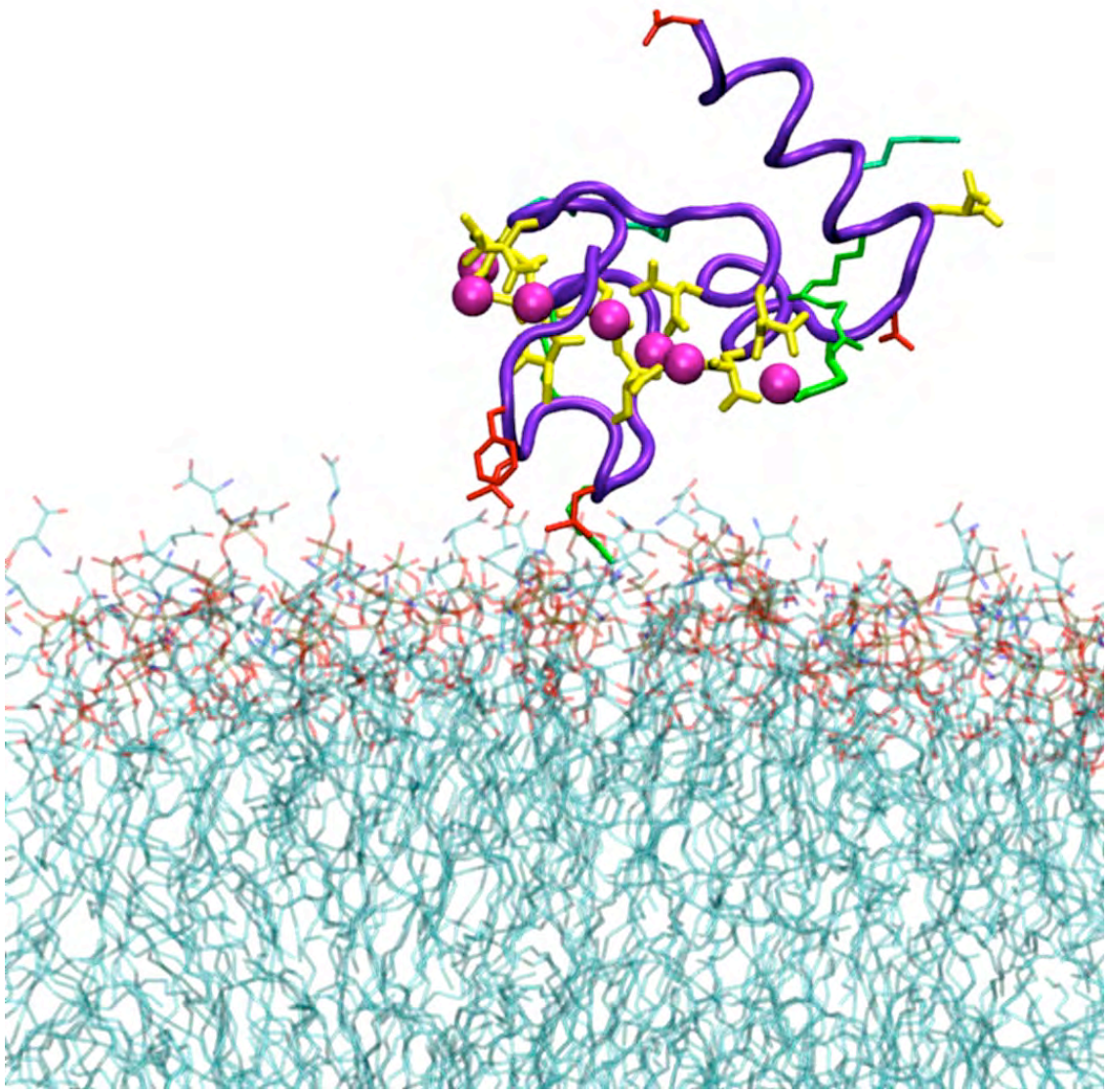


How do we construct an initial model for peripheral proteins in membranes?



Simulation of Binding with Full Membrane Representation





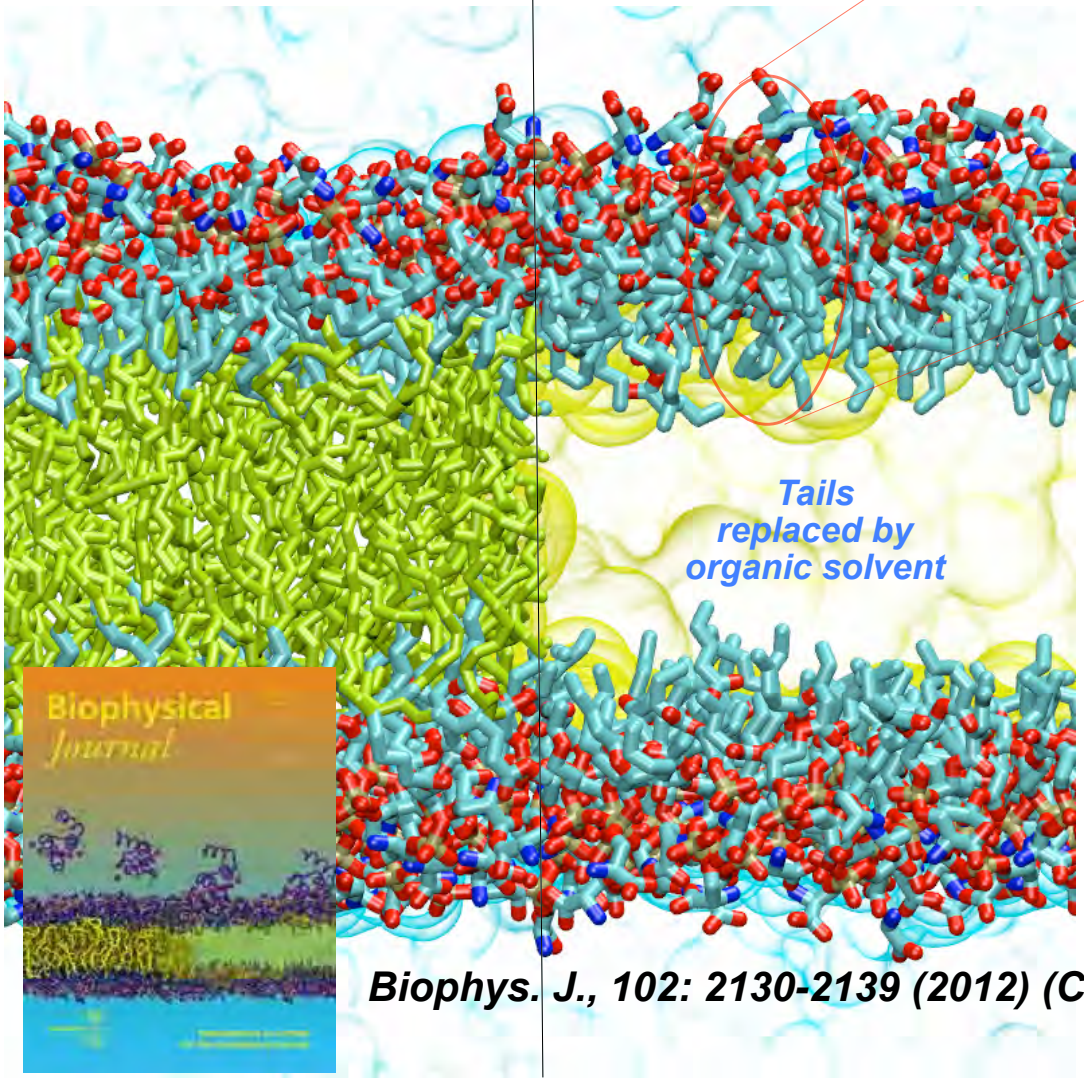
Partial list of technical problems:

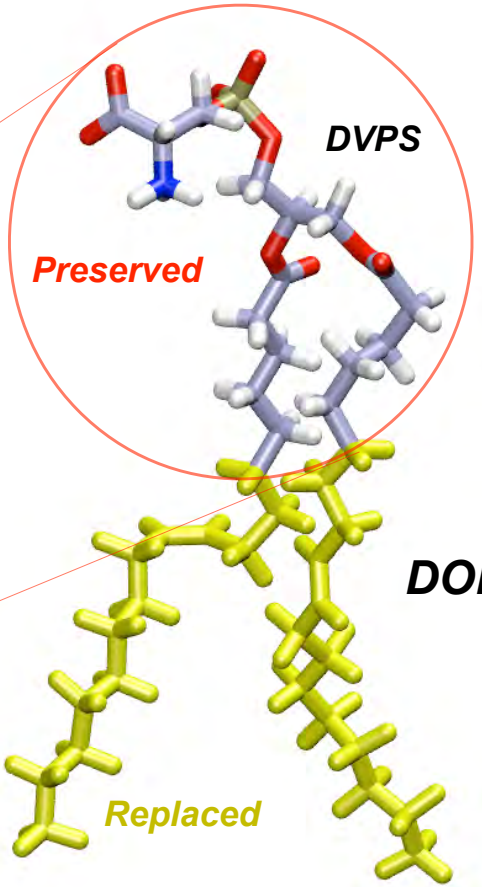
- Biased simulations
- Unknown depth of insertion
- Single binding event
- Frequently failing
- **Minimal lipid reorganization**

Z. Ohkubo and E. T., **Structure**, 16: 72-81 (2008)

HMMM model

Highly Mobile Membrane Mimetic model





Advantages

Increased mobility of lipids

Retain explicit headgroups allowing for atomic details

Biophys. J., 102: 2130-2139 (2012) (Cover Article)

Zenmei Ohkubo

Mark Arcario

Taras Pogorelov

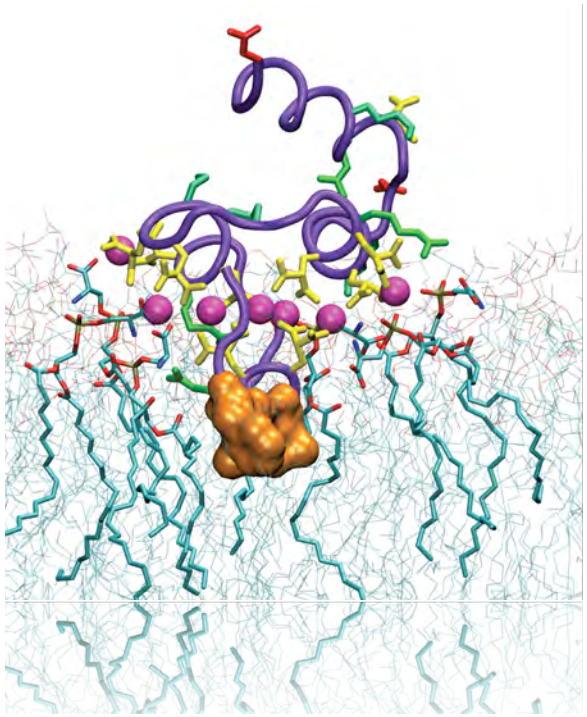
Josh Vermaas

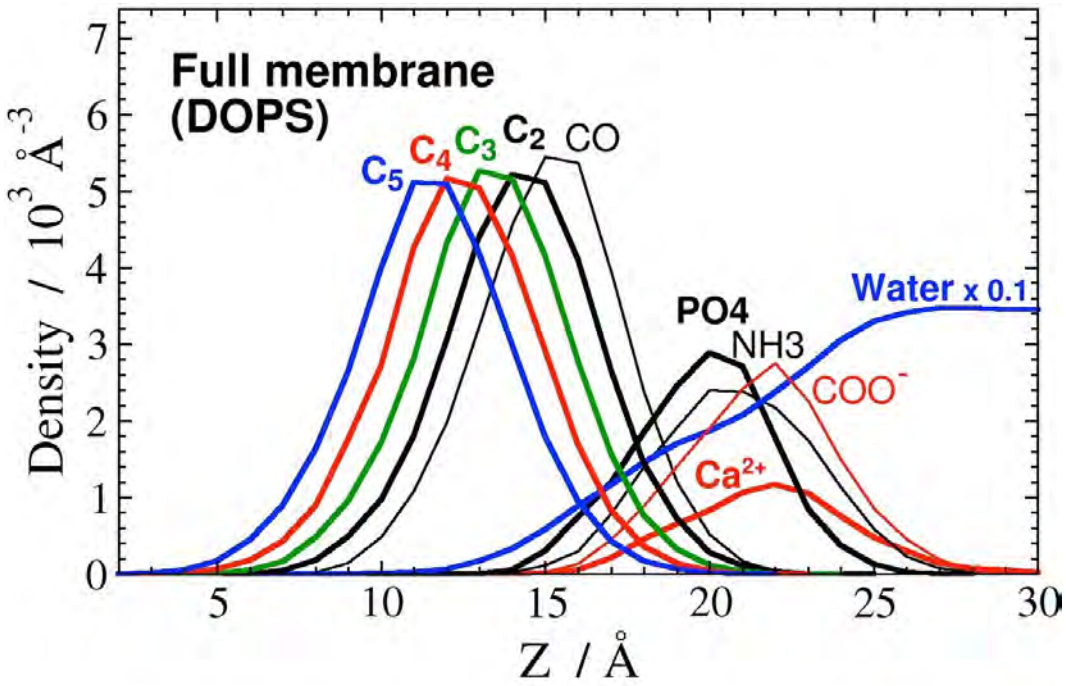
Javier Baylon

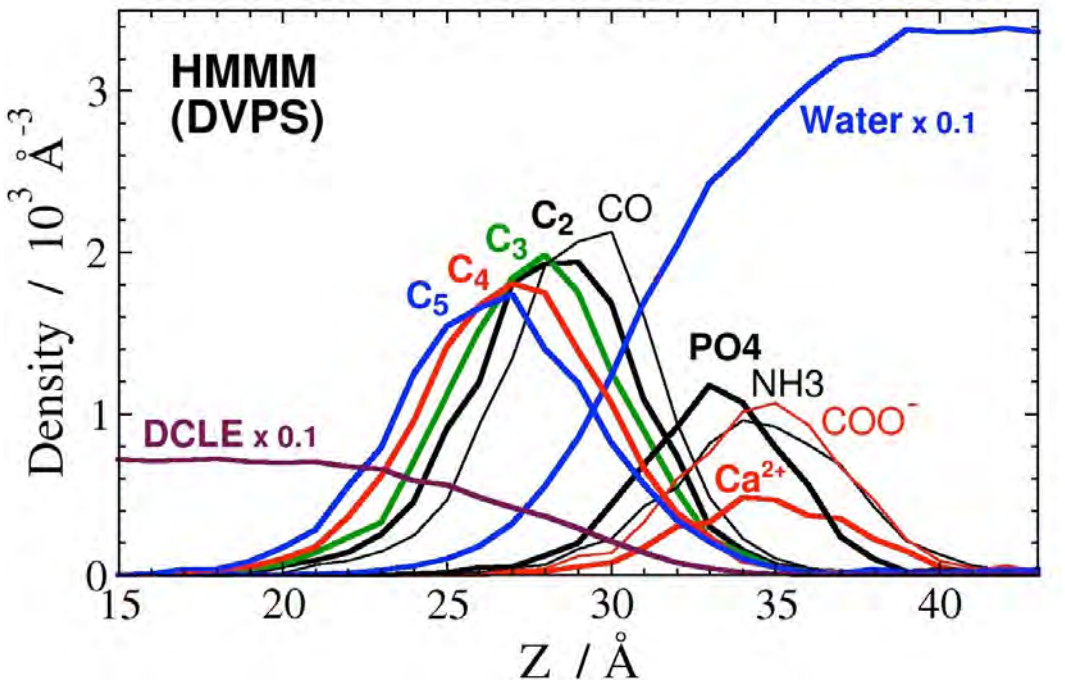
HMMM- Preserving the "Face" of the Lipid Bilayer

Perfect match in the membrane profile particularly in the head group region

Critical for proper description of lipid protein interactions

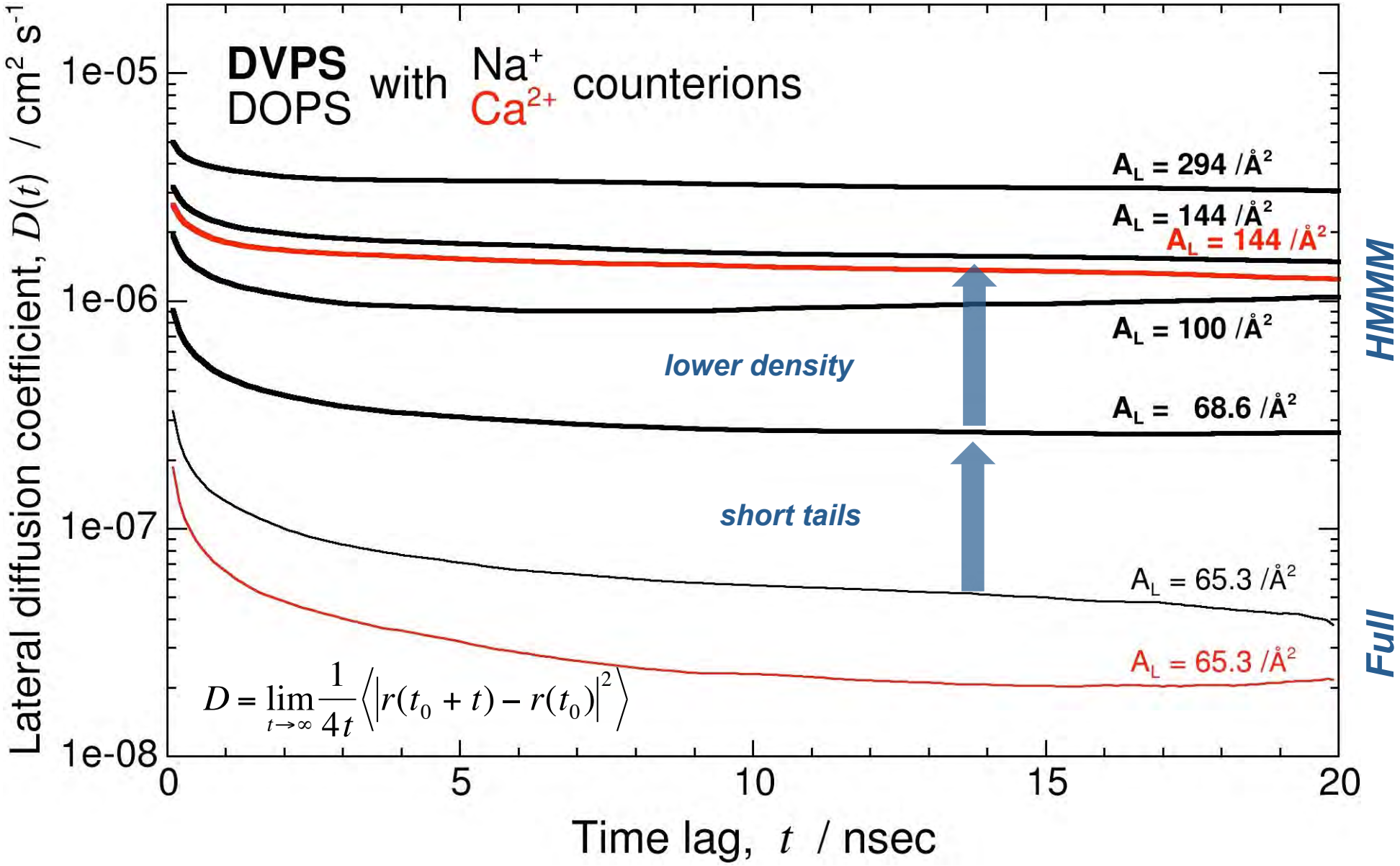






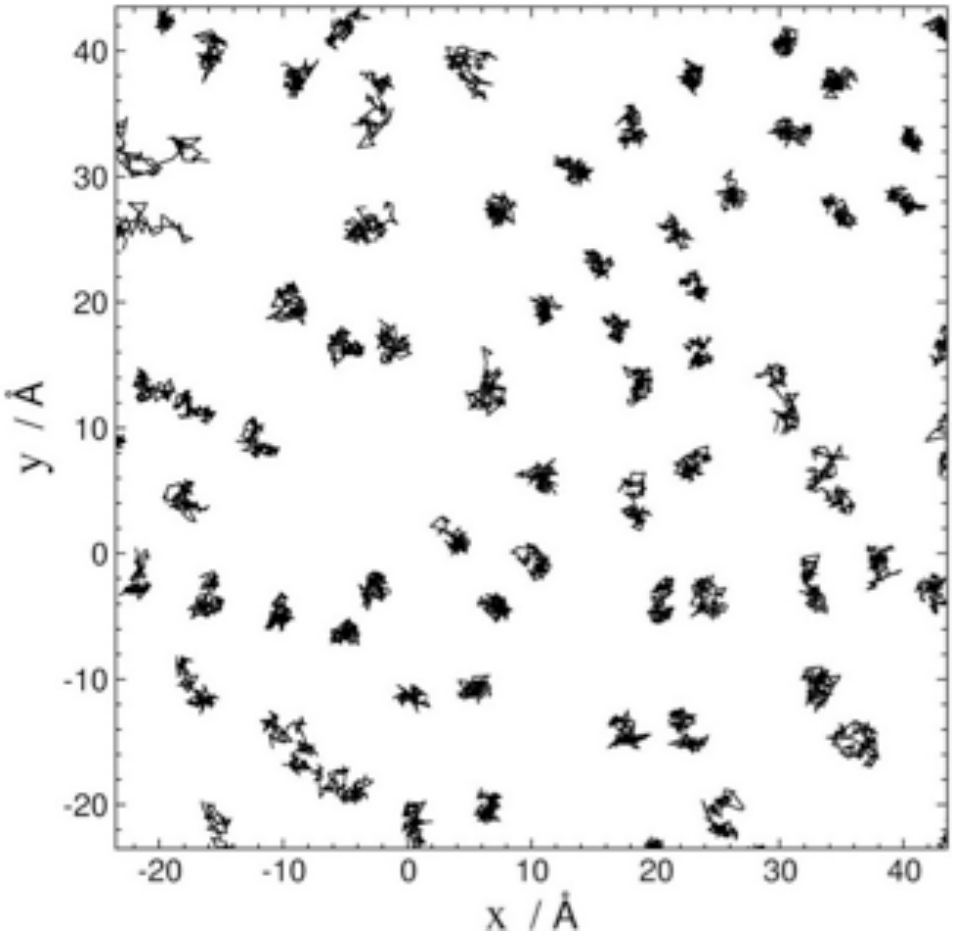
Enhanced Lipid Lateral Diffusion

Without Compromising Atomic Details of the Headgroups

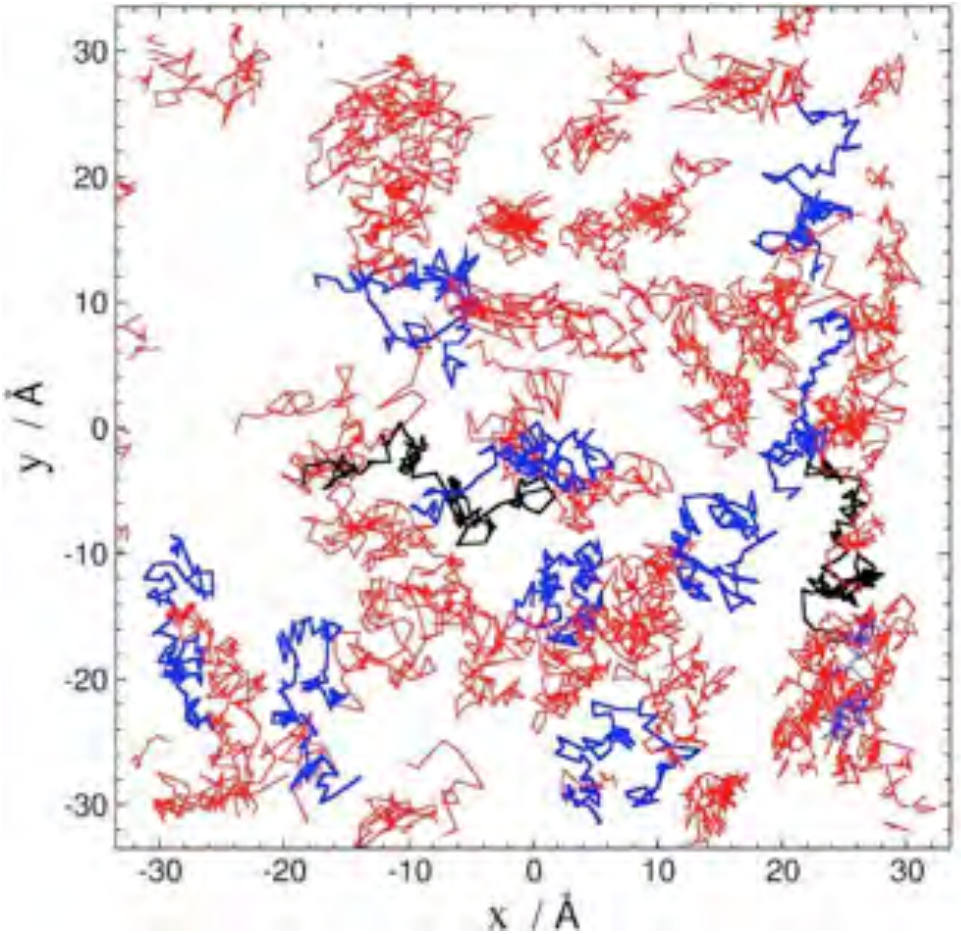


Enhanced Lipid Lateral Diffusion

Without Compromising Atomic Details of the Headgroups

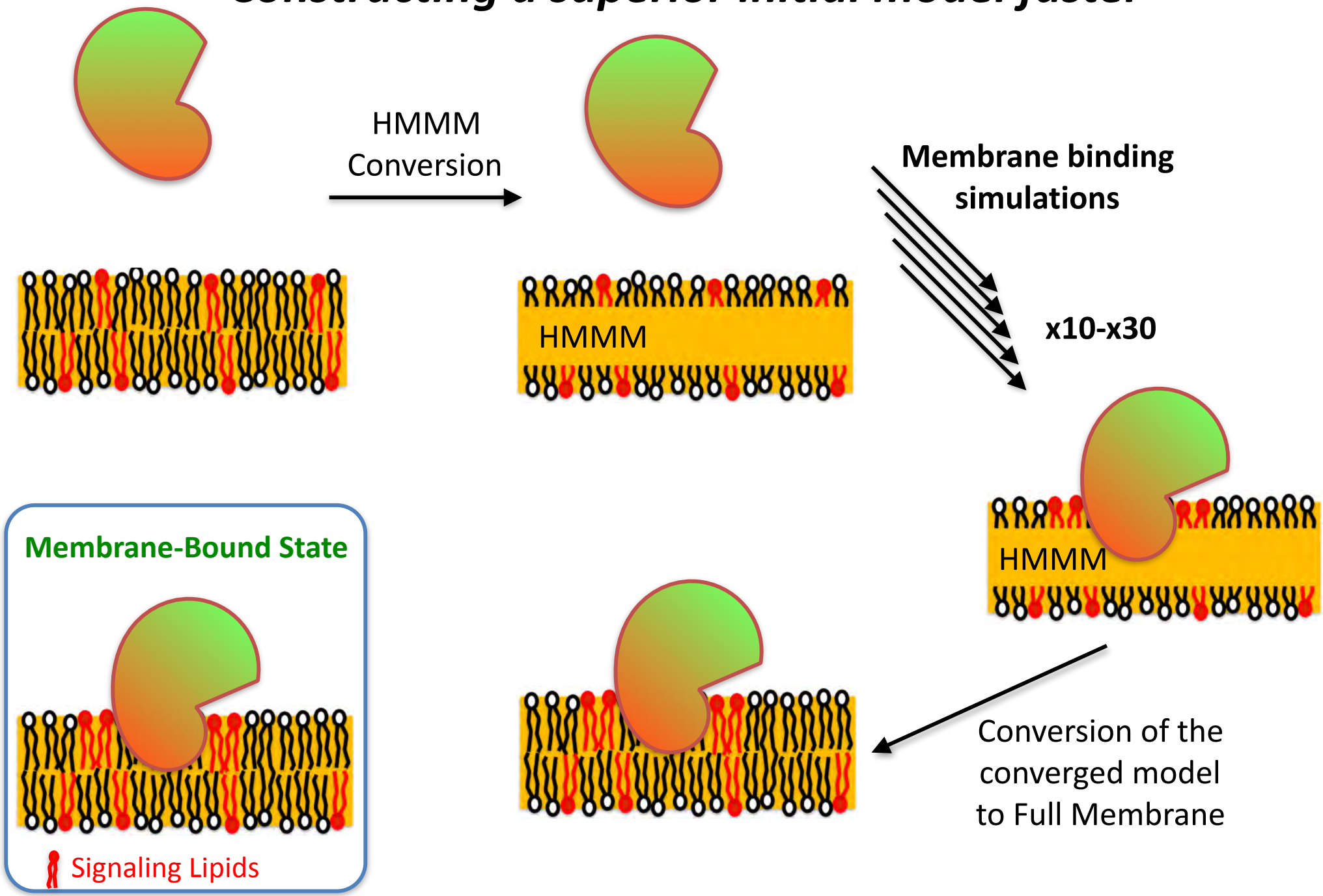


Conventional membrane (10 ns)



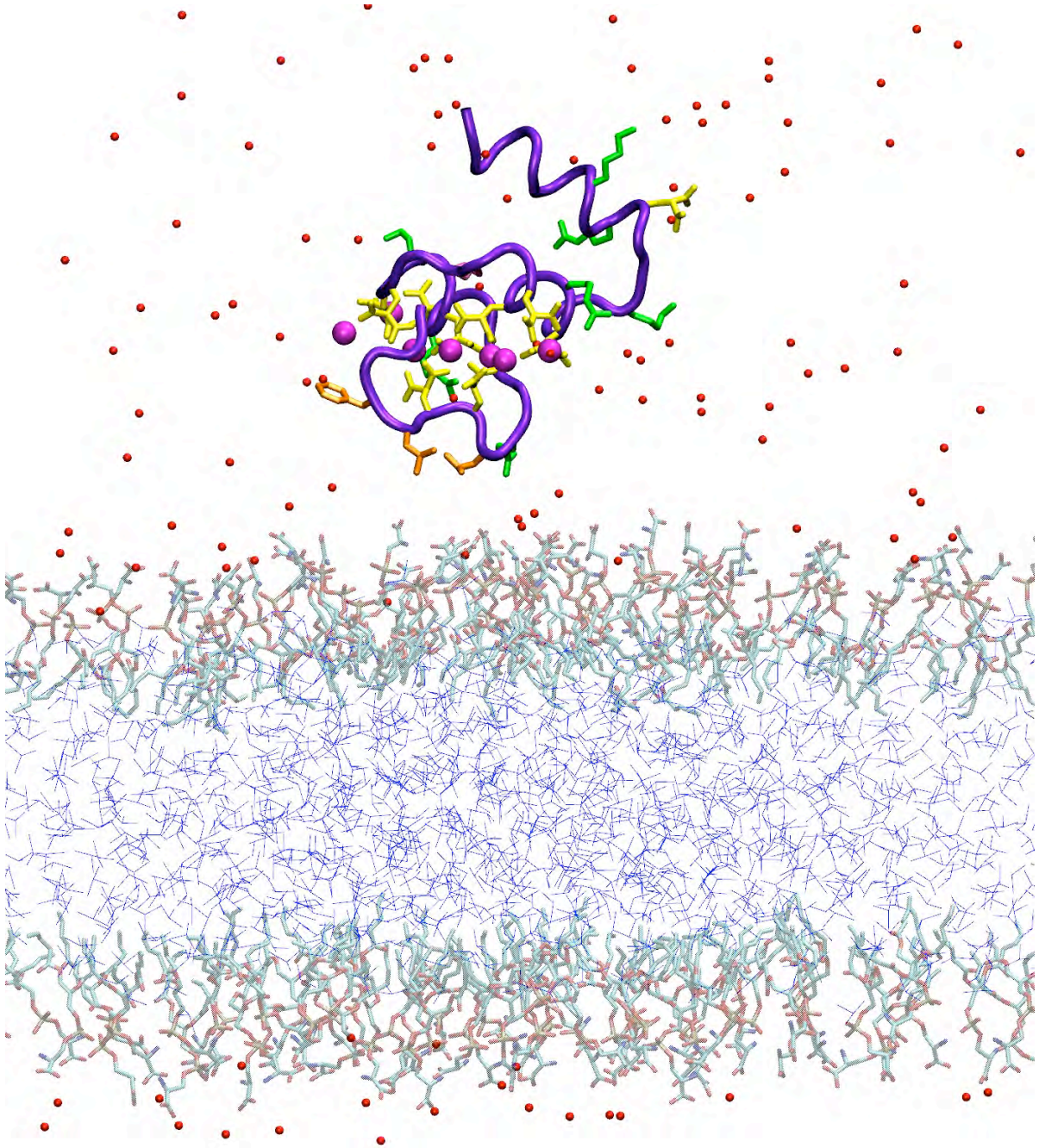
HMMM membrane (1 ns)

HMMM accelerated sampling of lipid-protein interactions
Constructing a superior initial model faster

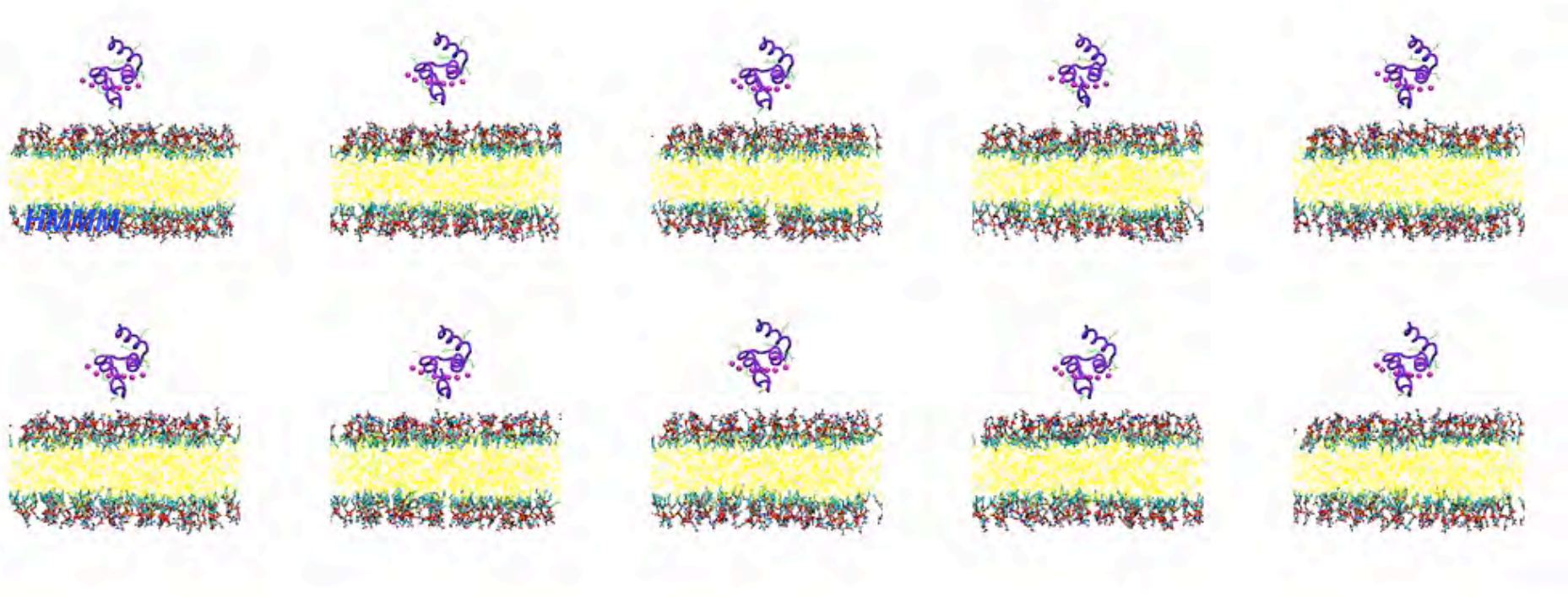


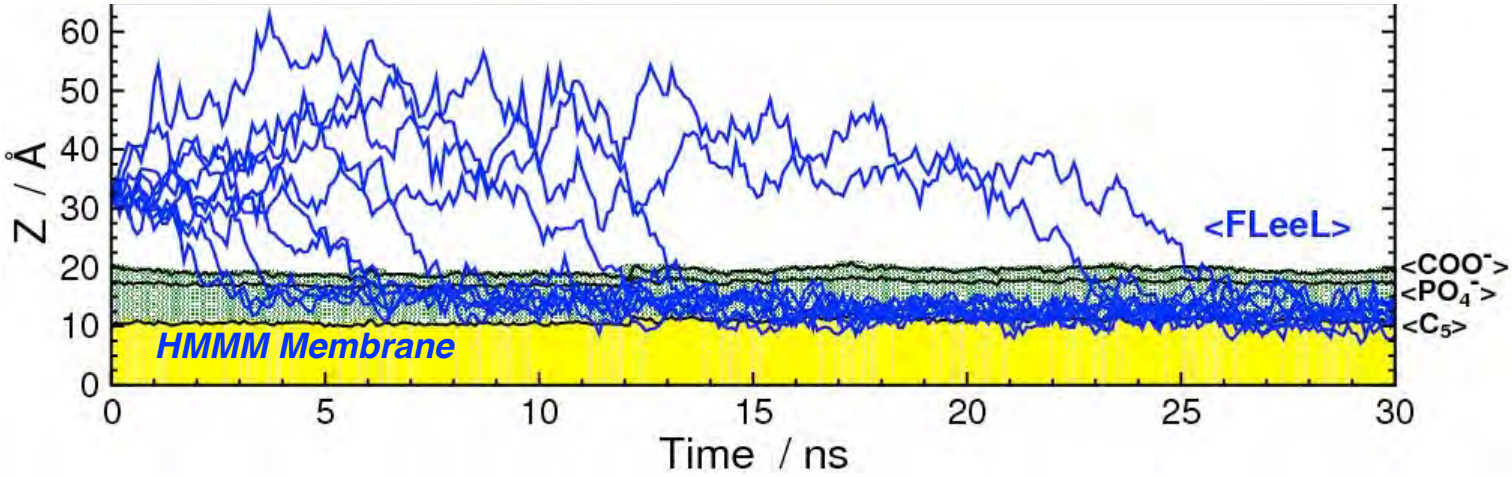
PS-Dependent Spontaneous Insertion of FVII-GLA

Zenmei Ohkubo



Spontaneous, Unbiased Membrane Binding Accelerated Process Allows for better sampling (***n = 10***)

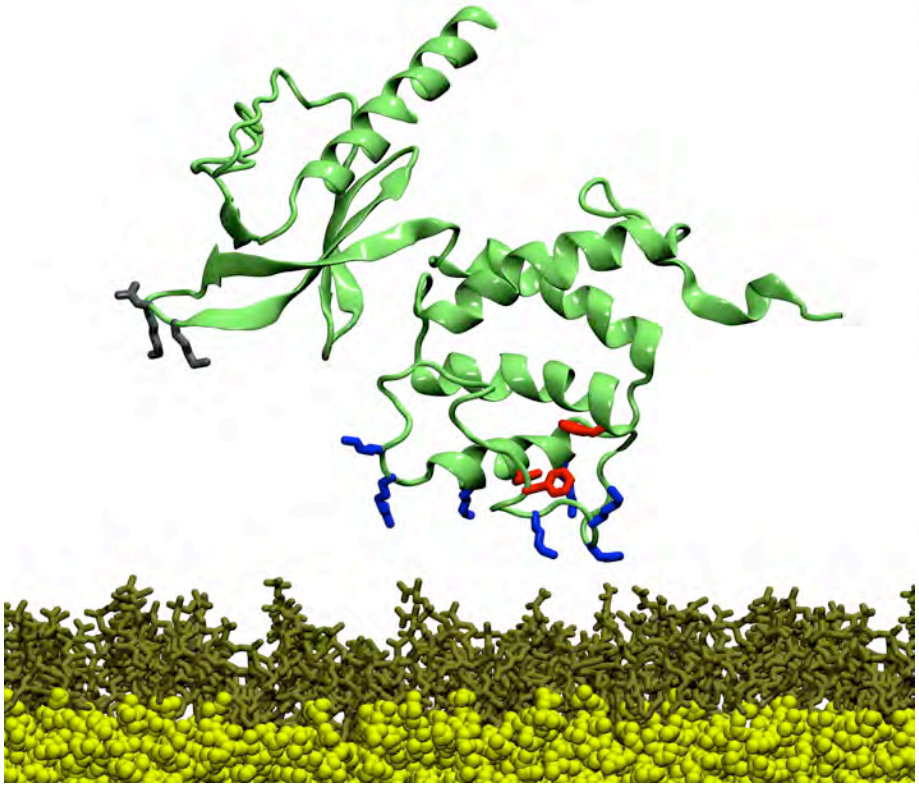




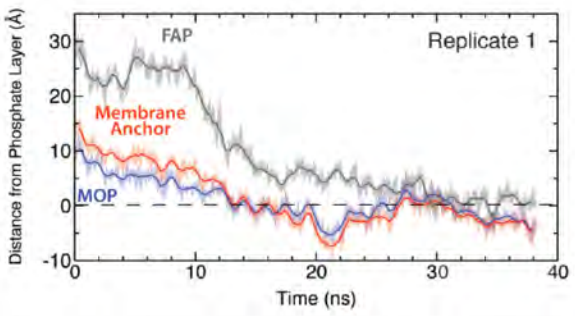
Z. Ohkubo, …, E.T., **Biophys. J.**, 102: 2130-2139 (2012) (Cover Article)

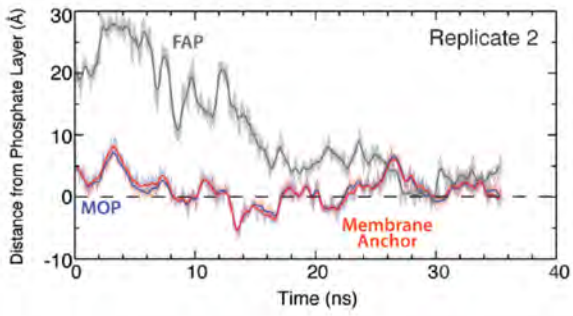
Mark Arcario

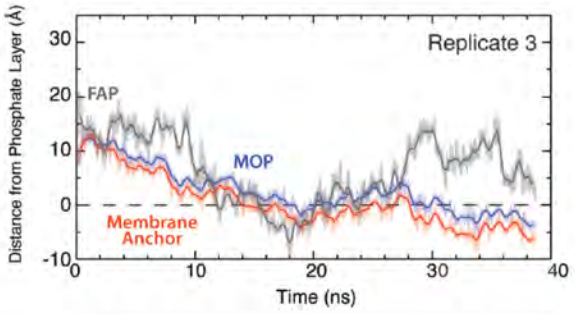
PS-Dependent Membrane Binding of Talin

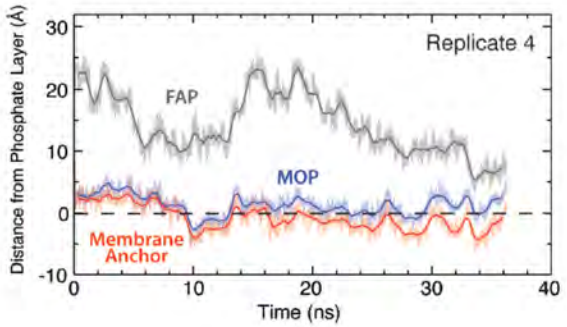


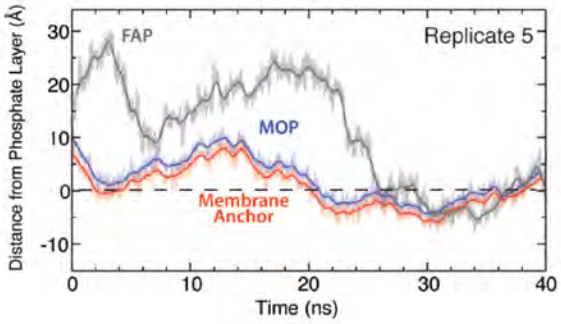
Five independent membrane binding simulations







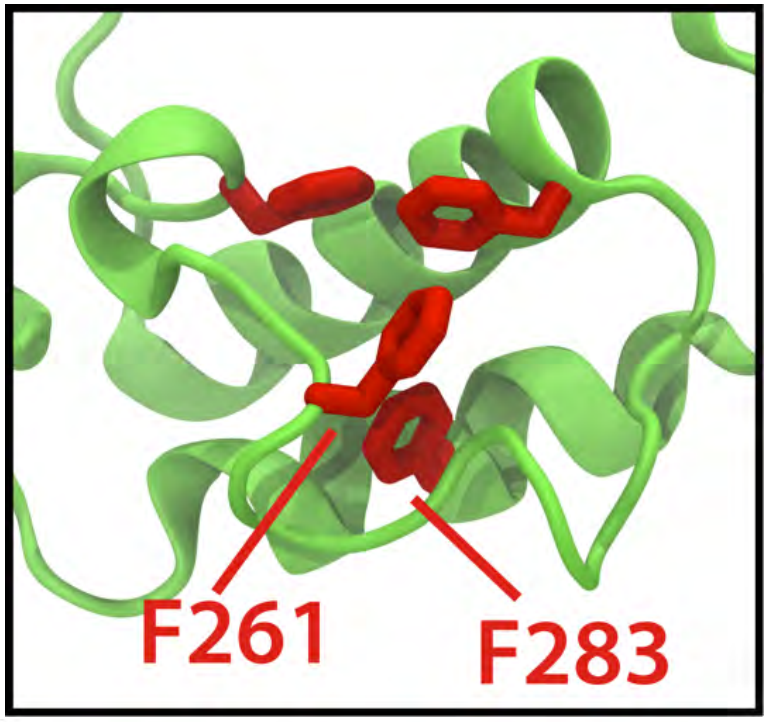




Final model converted to **full membrane**
Stable in 100 ns simulations

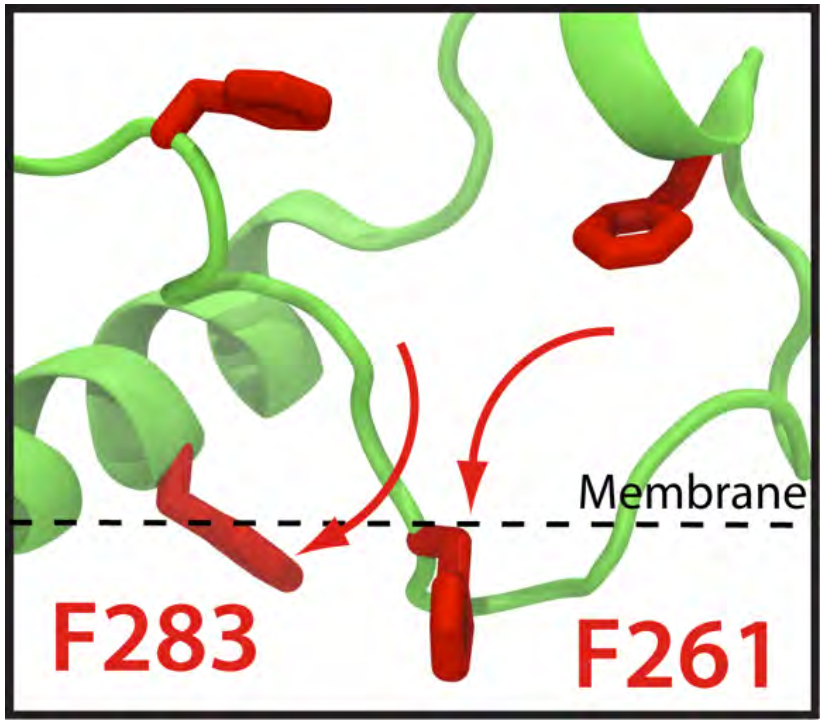
M. Arcario and ET, **Biophys. J.**, 107: 2059–2069 (2014).

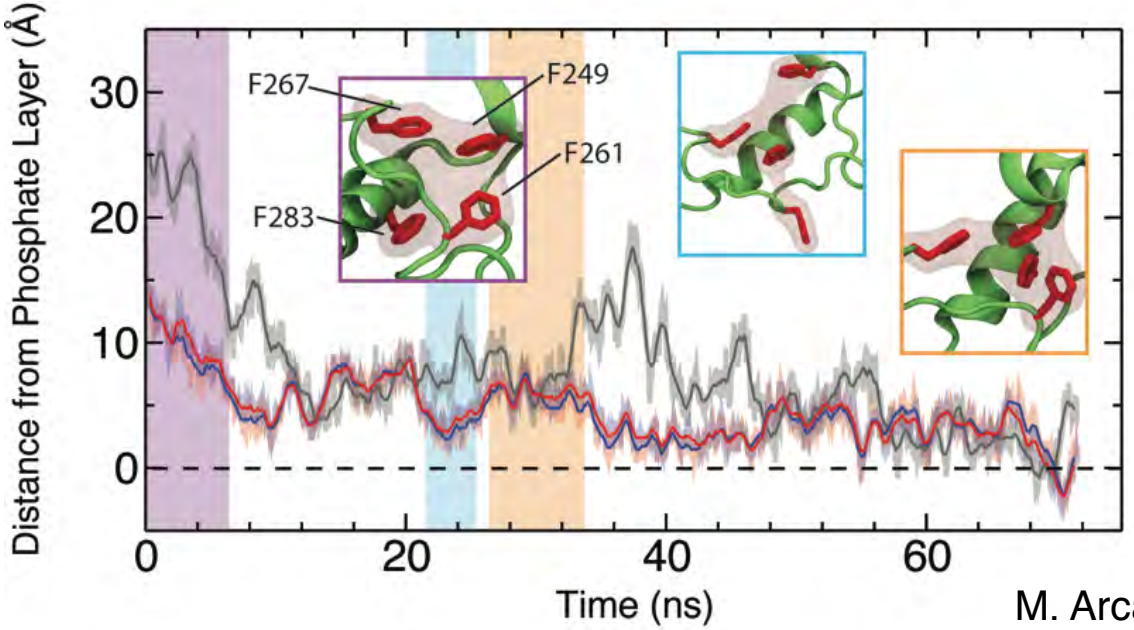
Revealing the *Hydrophobic Anchor*



20 ns

Membrane Binding Simulation

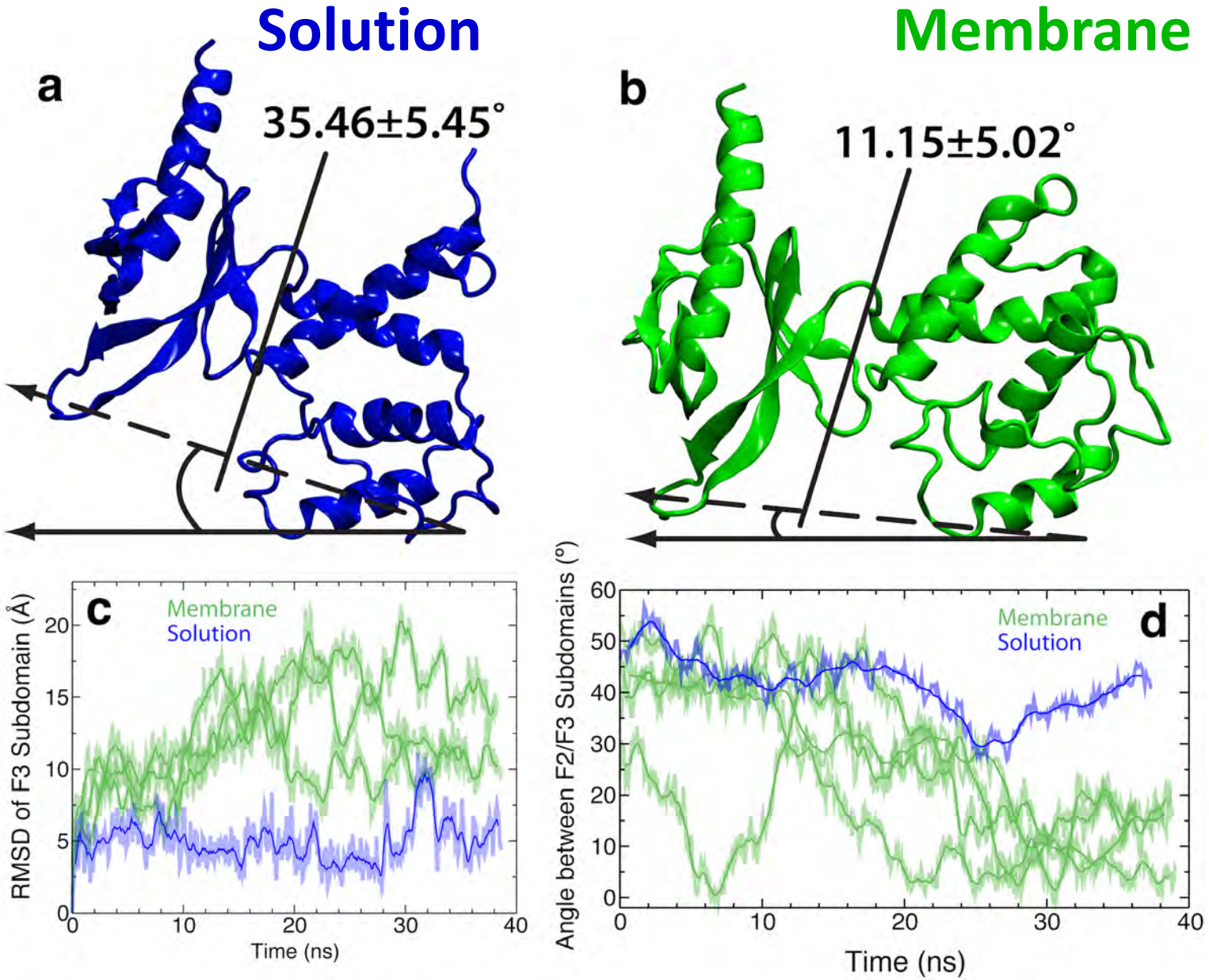




- Snorkeling of lysine acts as a switch which releases a conserved phenylalanine anchor (F261 & F283) into the membrane
- Reformation of the hydrophobic pocket causes looser binding of talin; suggests a mechanism for unbinding of protein

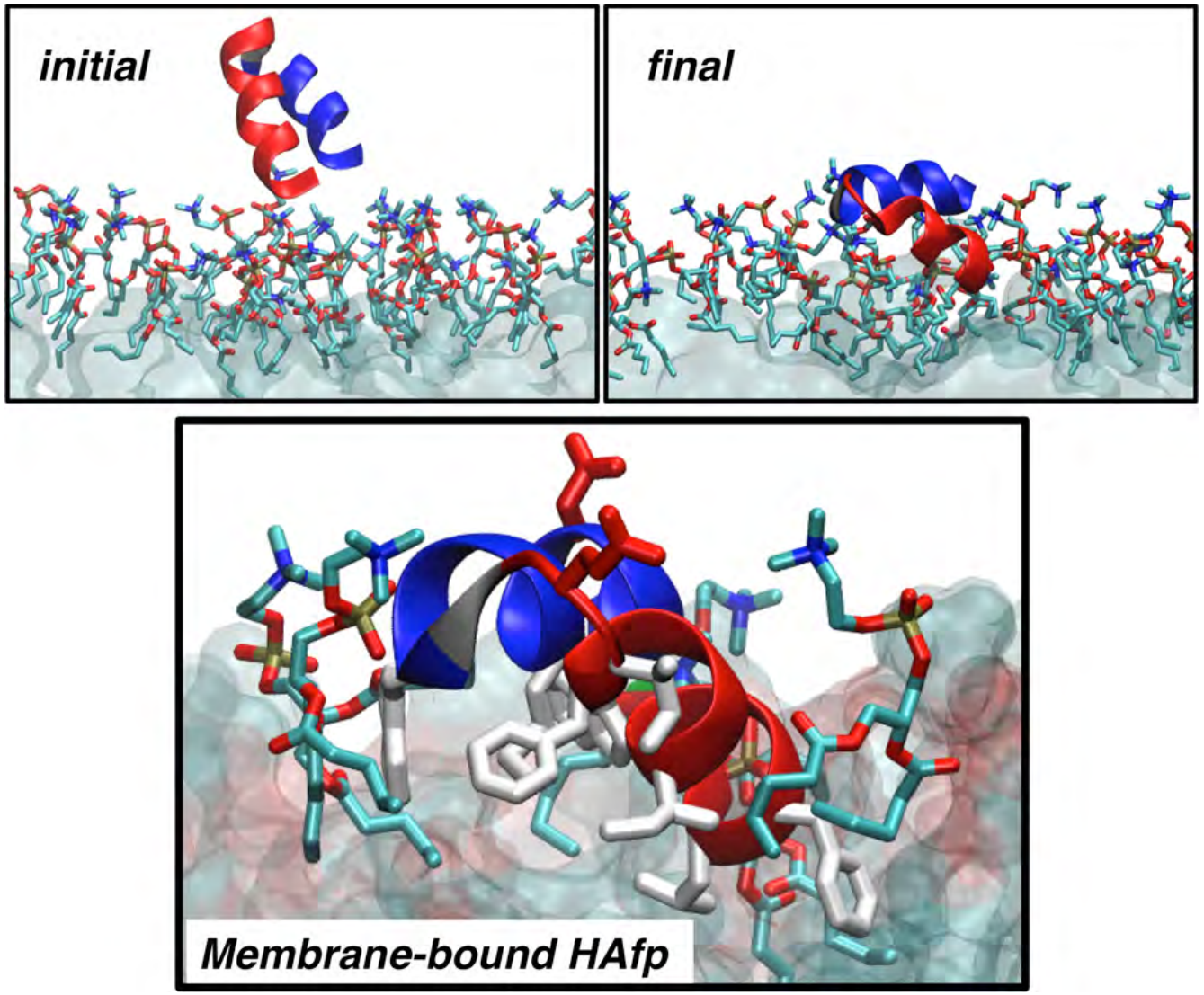
M. Arcario and ET, **Biophys. J.**, 107: 2059–2069 (2014).

Membrane Induced Domain Rearrangement of Talin



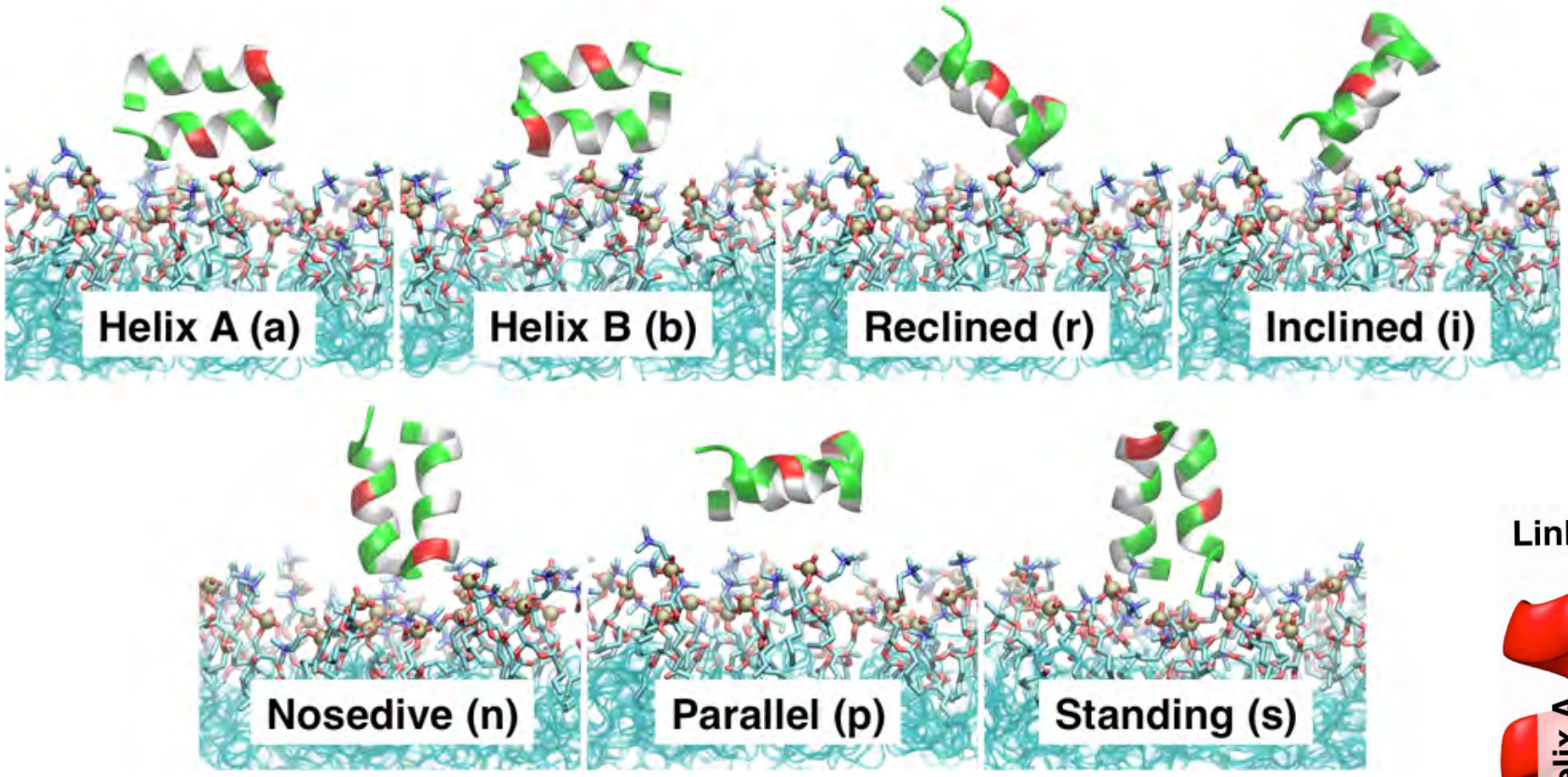
M. Arcario and ET, **Biophys. J.**, 107: 2059–2069 (2014).

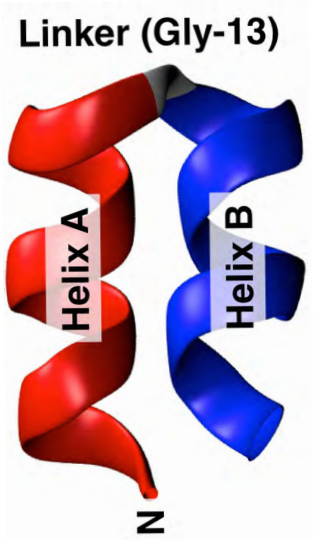
Membrane Binding of Influenza Hemagglutinin Fusion Peptide



J. Baylon and E. T., **J. Phys. Chem.B**, 2015, in press.

Membrane Binding of Influenza Hemagglutinin Fusion Peptide



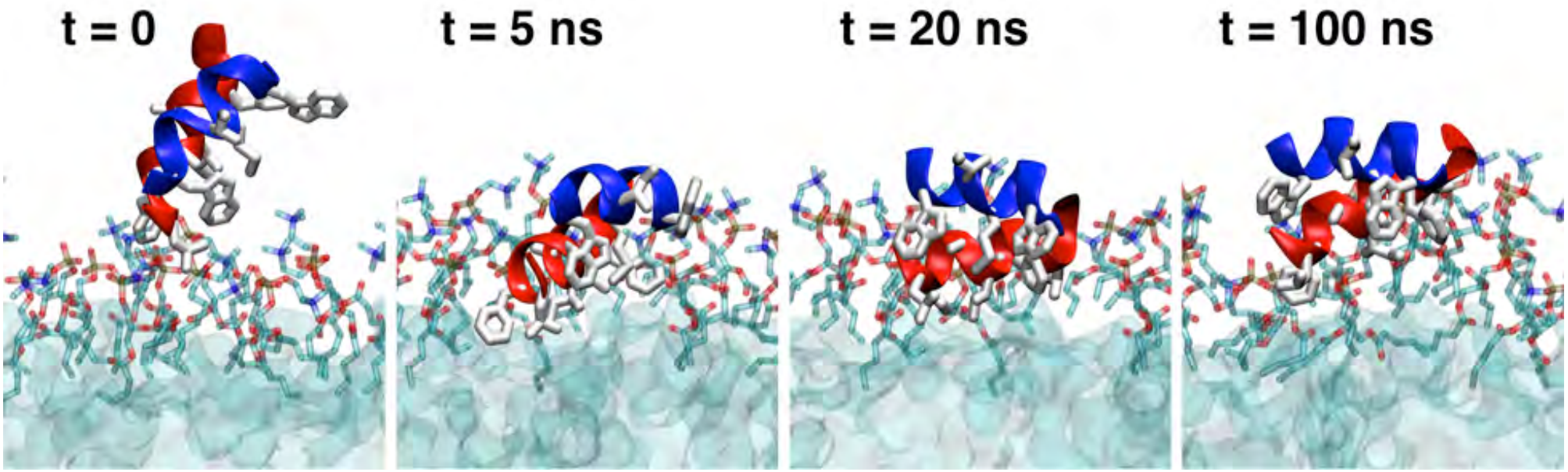


7 different initial orientation
each simulated 3 times

J. Baylon and E. T., **J. Phys. Chem.B**, 2015.

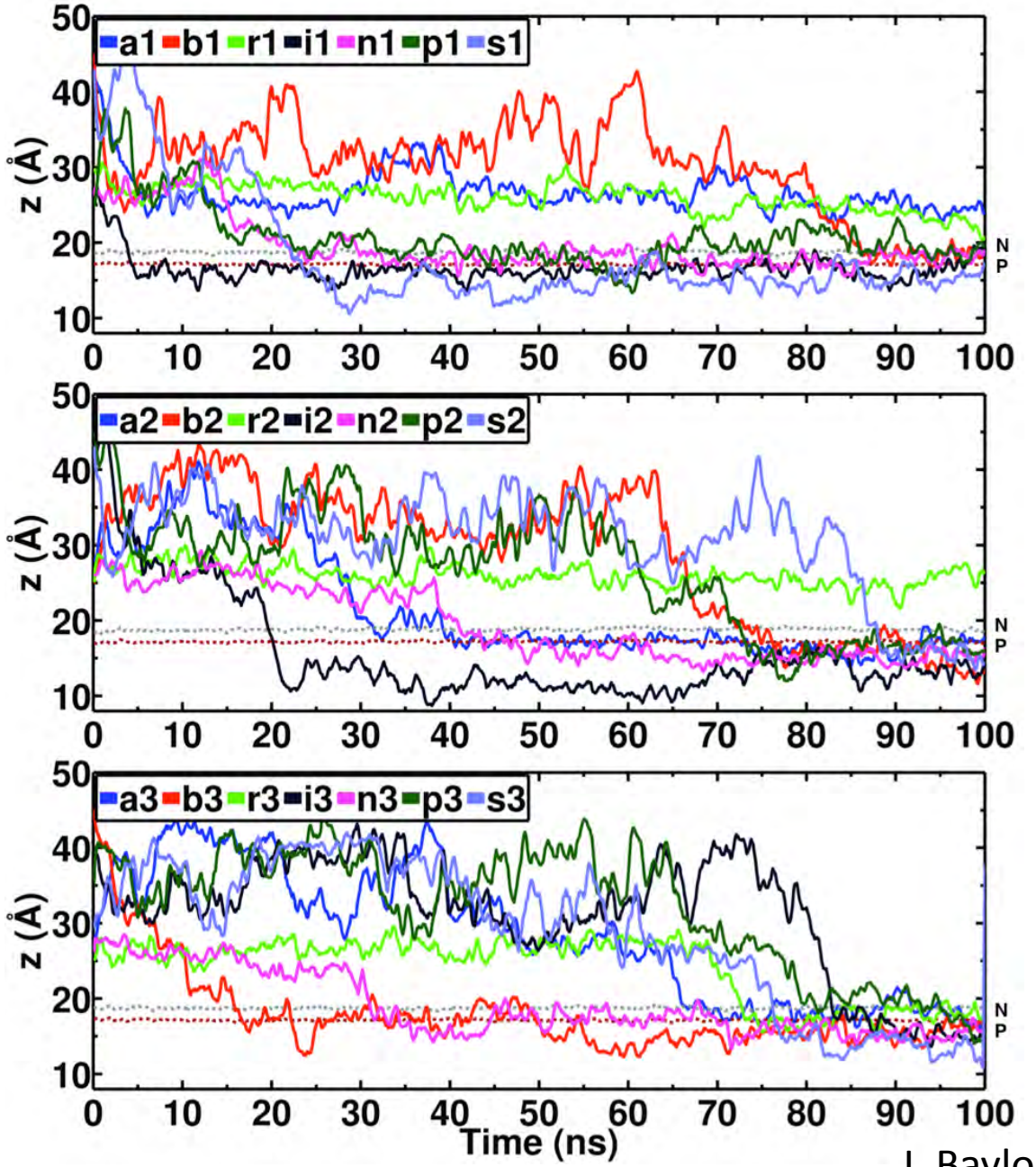
Membrane Binding of Influenza Hemagglutinin Fusion Peptide

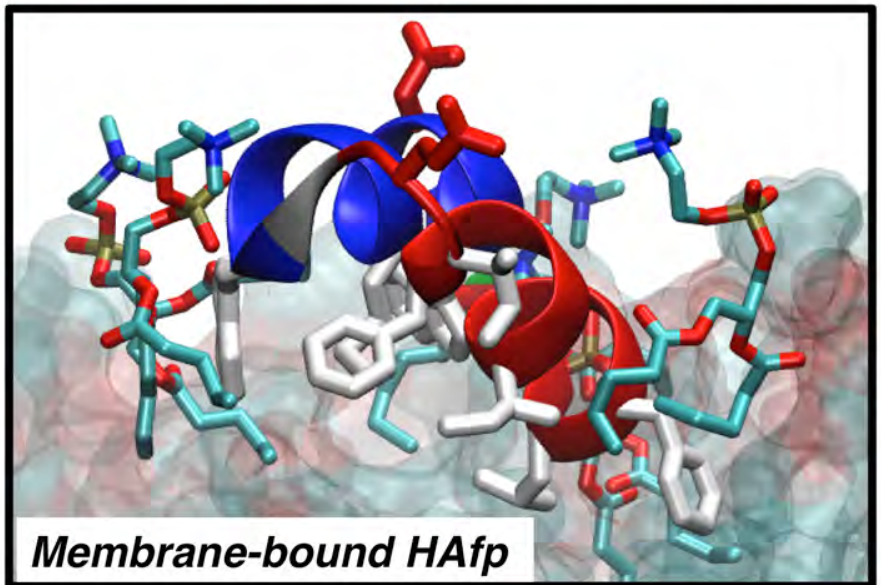
Spontaneous binding observed in the majority of the simulations:
21 independent simulations starting from 7 different orientations



J. Baylon and E. T., **J. Phys. Chem.B**, 2015.

Remarkable convergence of membrane binding simulations

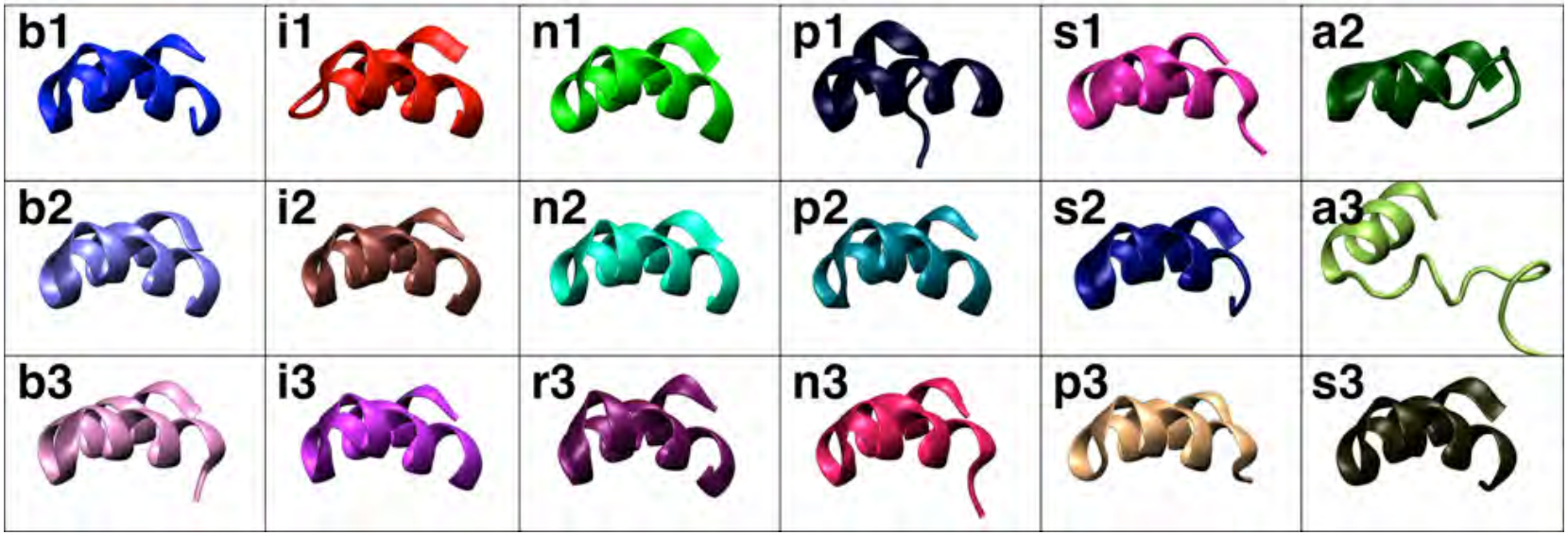




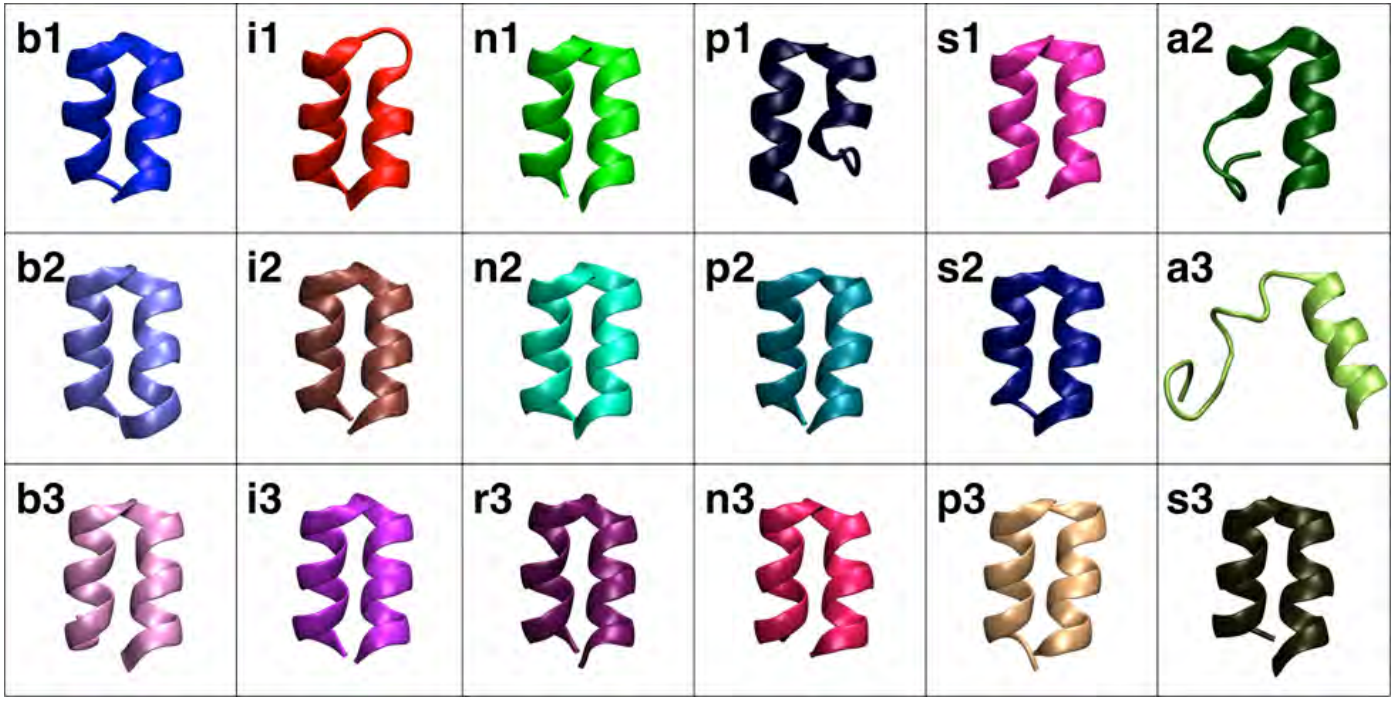
J. Baylon and E. T., **J. Phys. Chem.B**, 2015, in press.

Remarkable convergence of membrane binding simulations

Side View



Top View

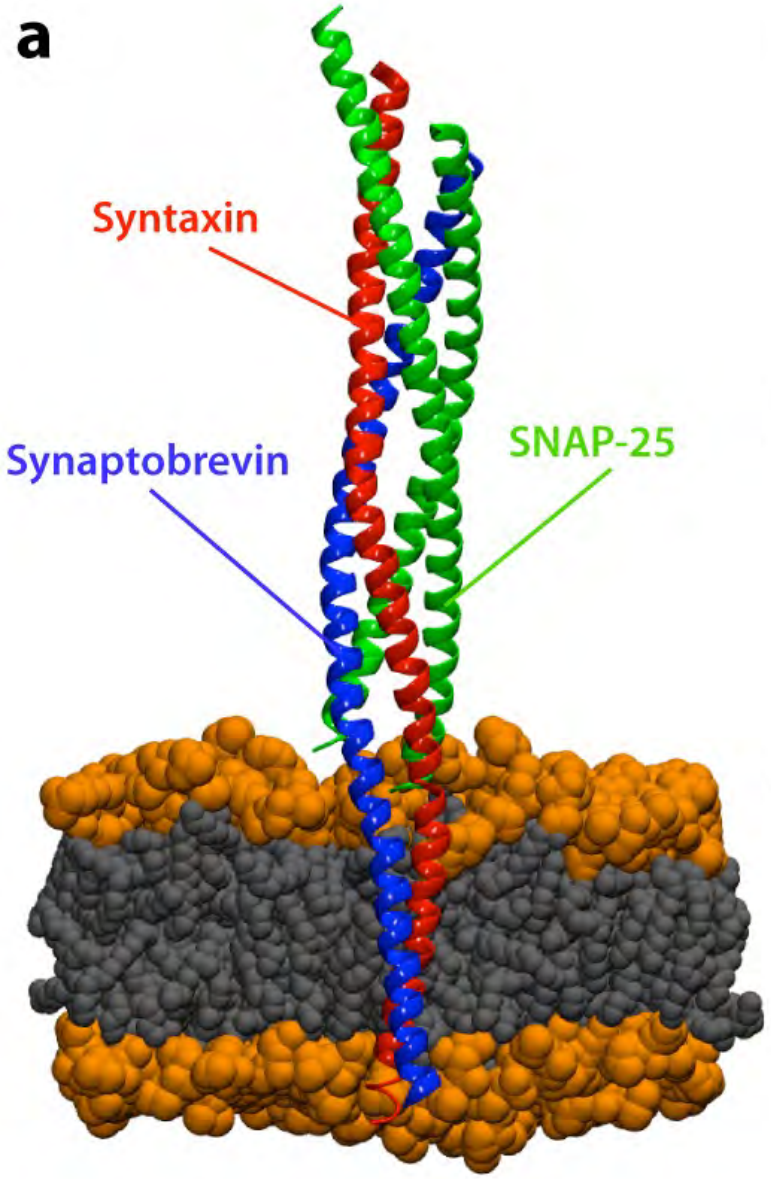


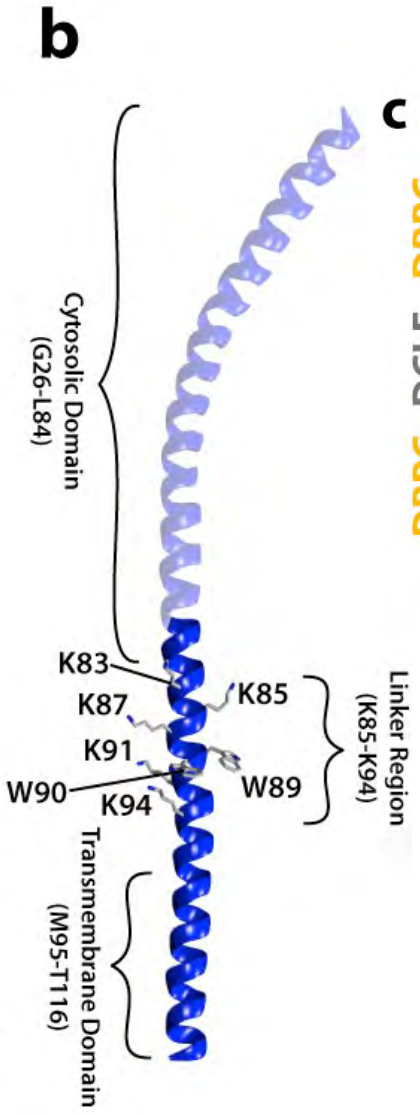
J. Baylon and E. T., **J. Phys. Chem.B**, 2015.

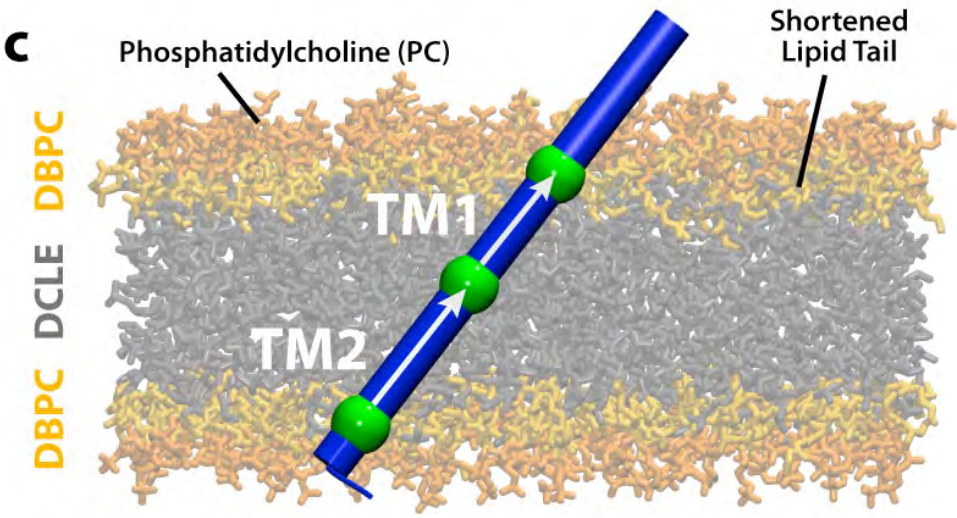
Robust Tilting of the Anchor Domain in Snare Protein Synaptobrevin


Mark Arcario

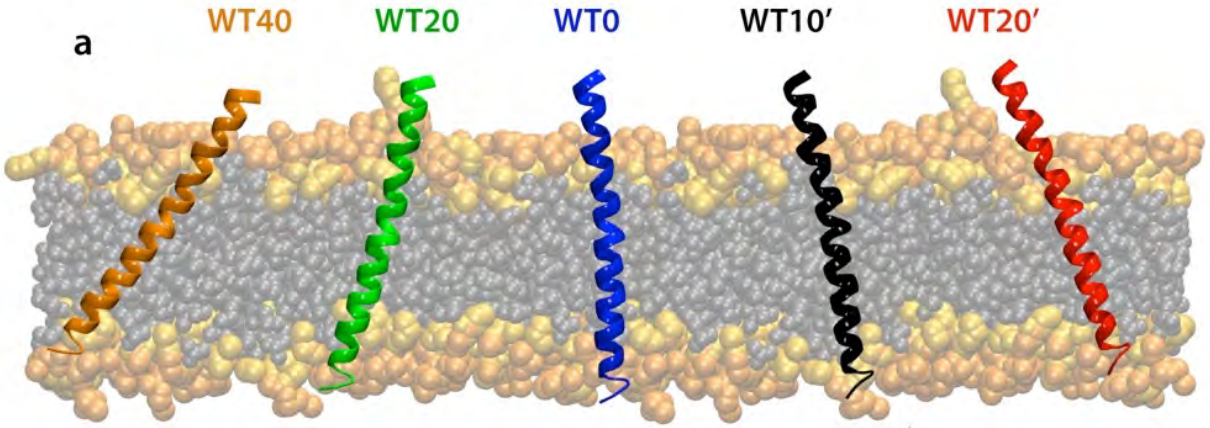

Andrew Blanchard



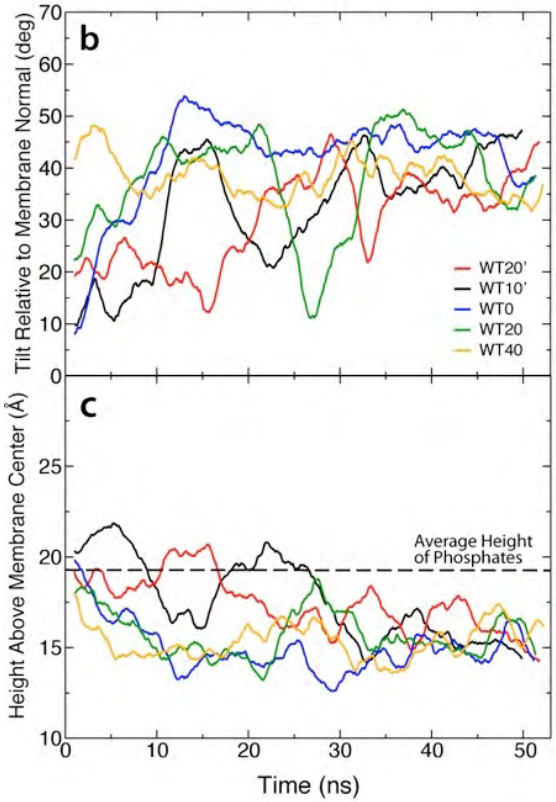


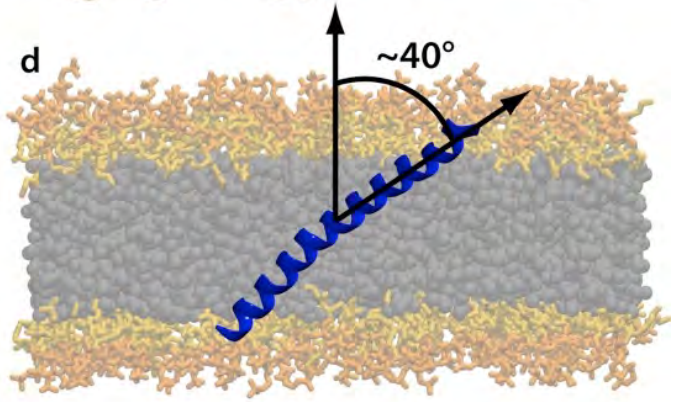


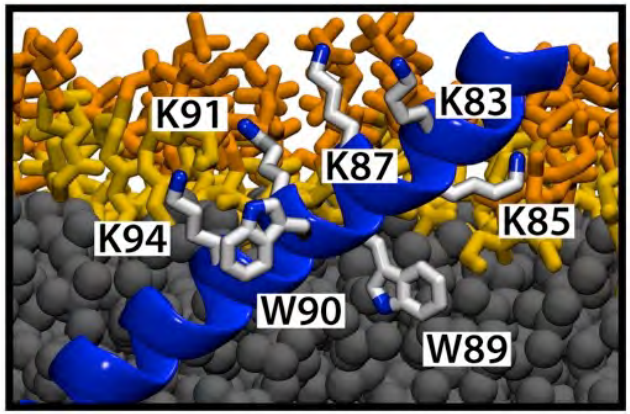
Robust Tilt Observed in Synaptobrevin



Membrane thickness mildly restrained:
one carbon/tail
k = 0.05 kcal/mole/A2

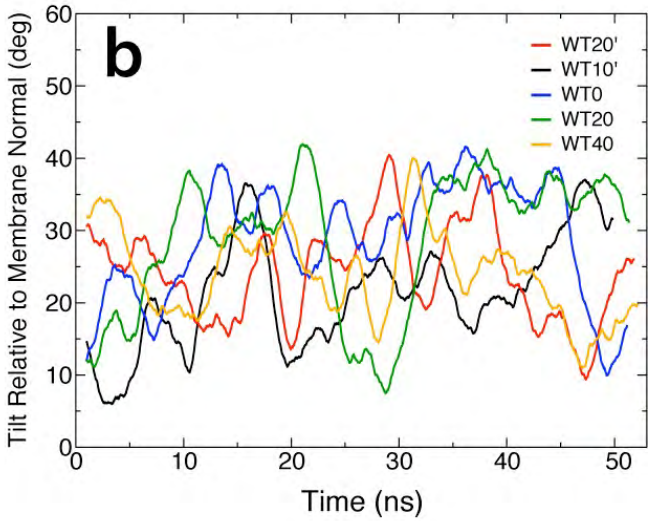


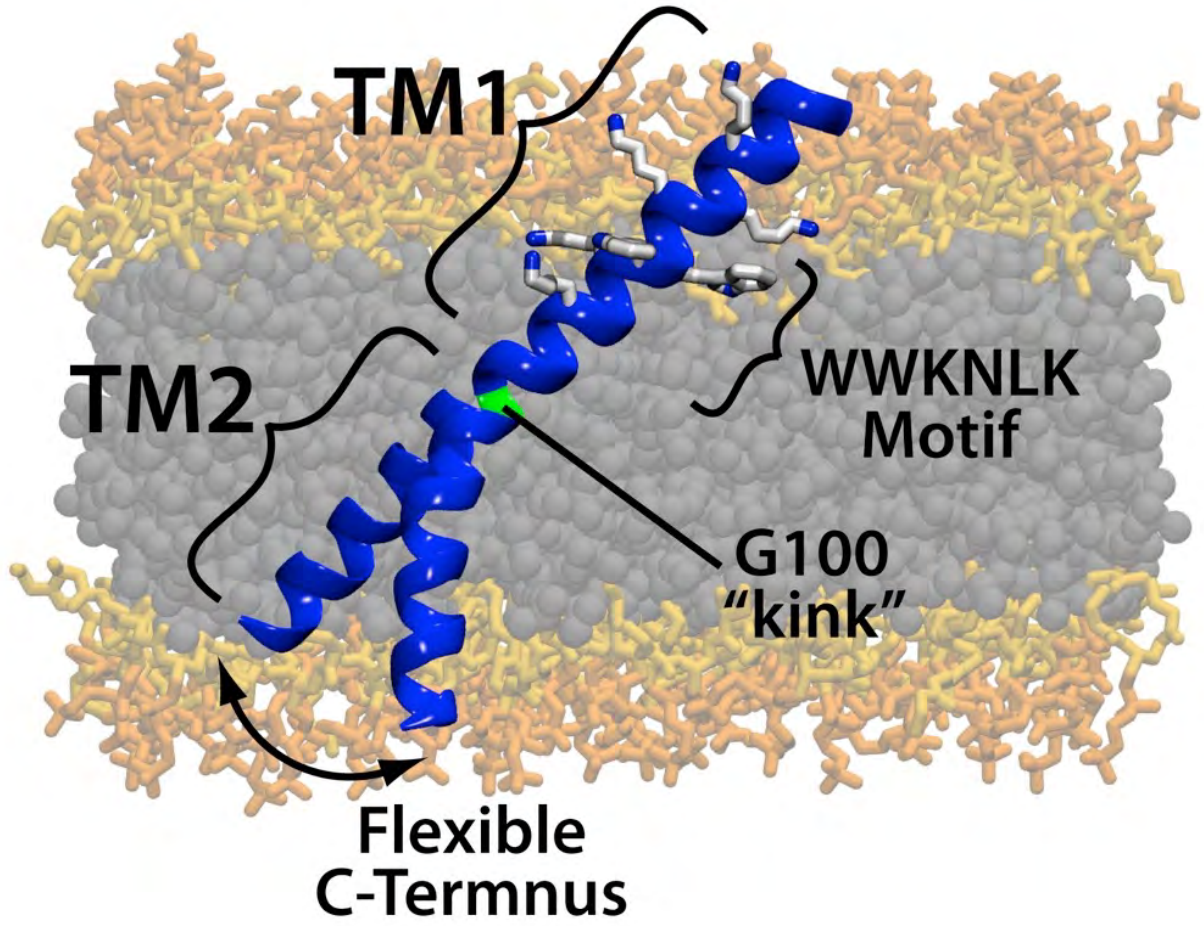




A. Blanchard*, M. Arcario*, K. Schulten, and ET, **Biophys. J.**, 107: 2112–21 (2014)

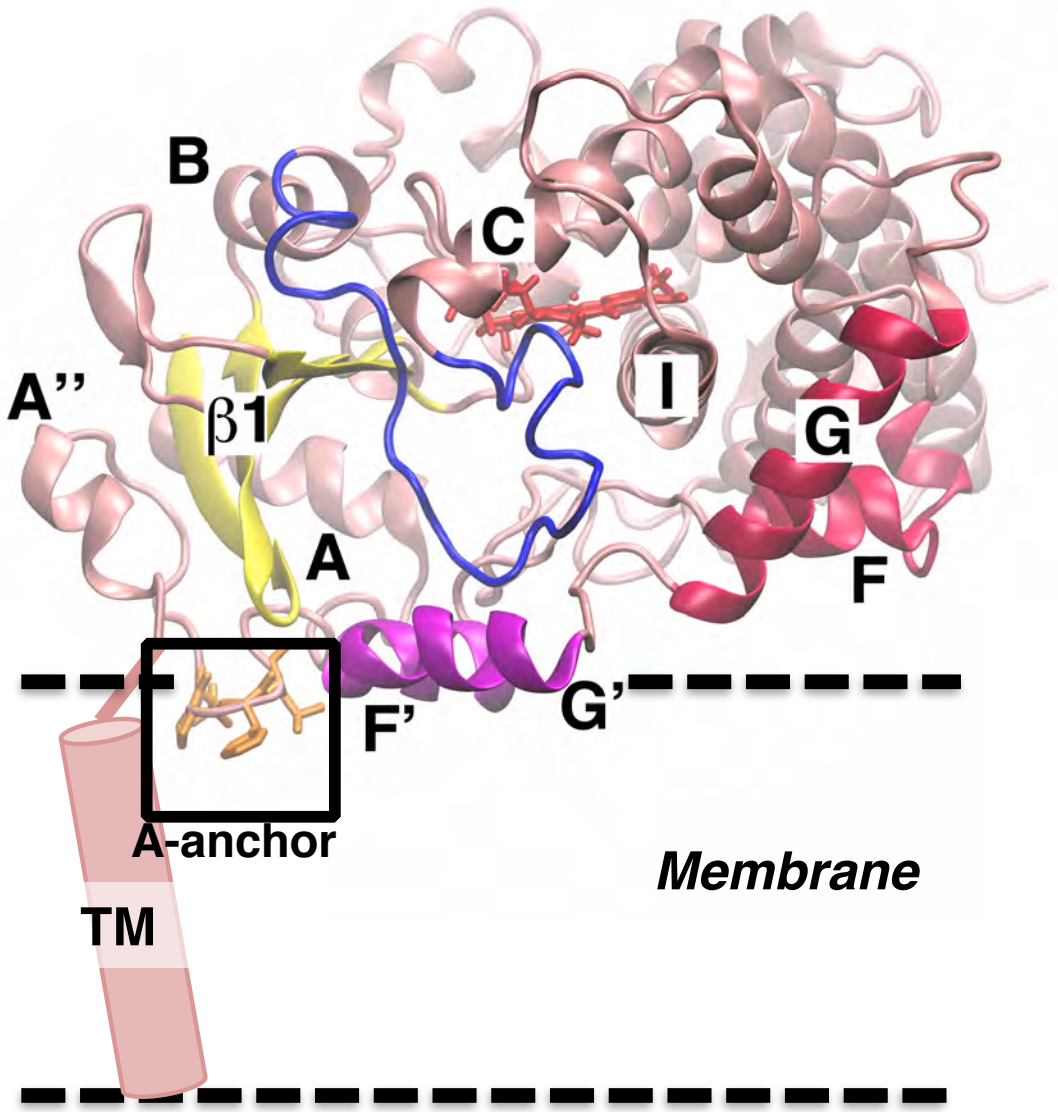
Identifying a Hinge





A. Blanchard*, M. Arcario*, K. Schulten, and ET, **Biophys. J.**, 107: 2112–21 (2014)

Cytochrome P450 3A4 (CYP3A4)

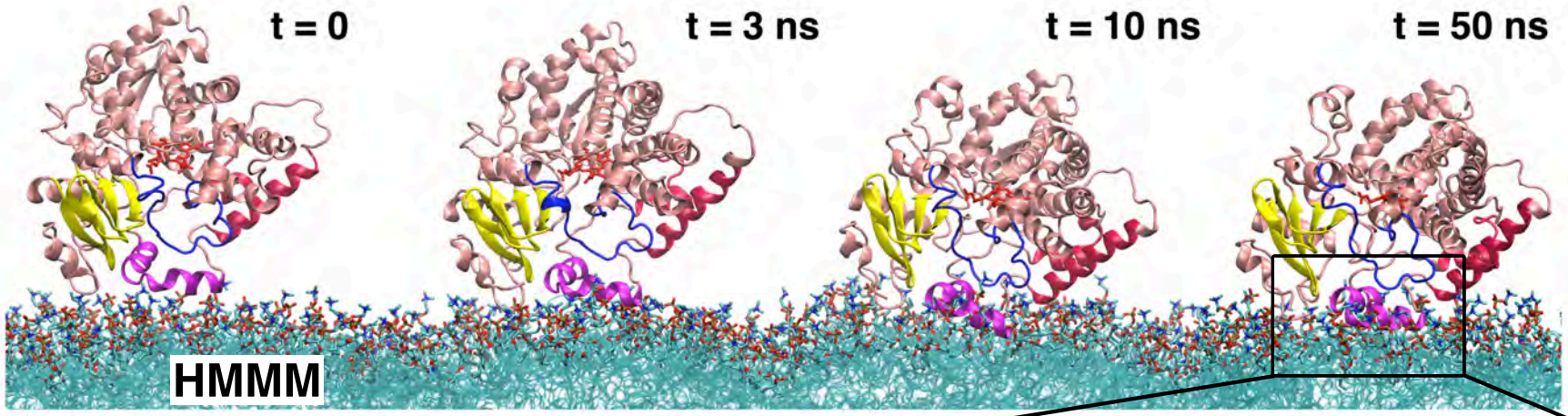


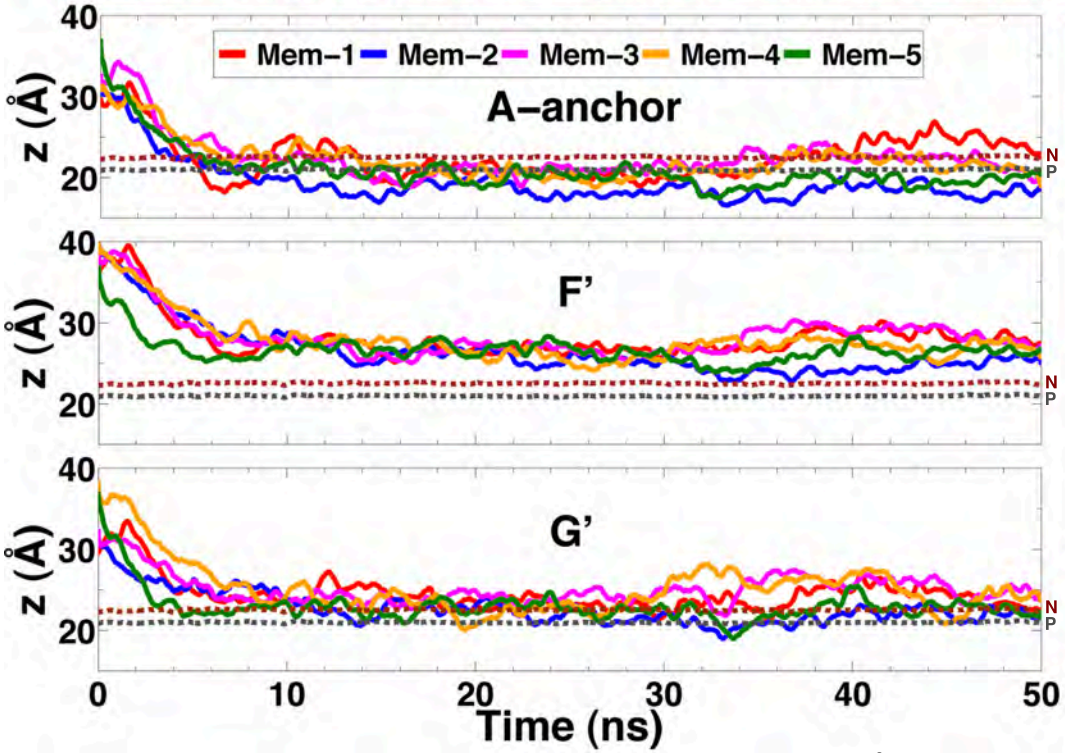
Yano *et al.*, J Biol Chem, 279: 38091-38094, 2004

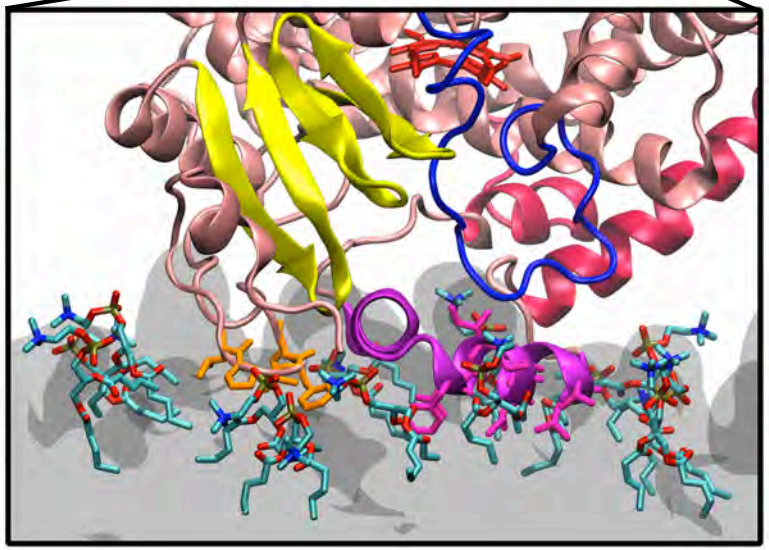
- Enzymes essential for the metabolism of xenobiotics and other compounds, found in all domains of life.
- In the human body, CYPs are membrane-bound proteins.
- The interaction with membrane mediates binding of substrates.
- **CYP3A4**: most abundant CYP in the human body, metabolizes about 50%- 60% of drugs that are metabolized in the body.

Insertion and Membrane-Induced Conformational Change of Cytochrome P450

Javier Baylon



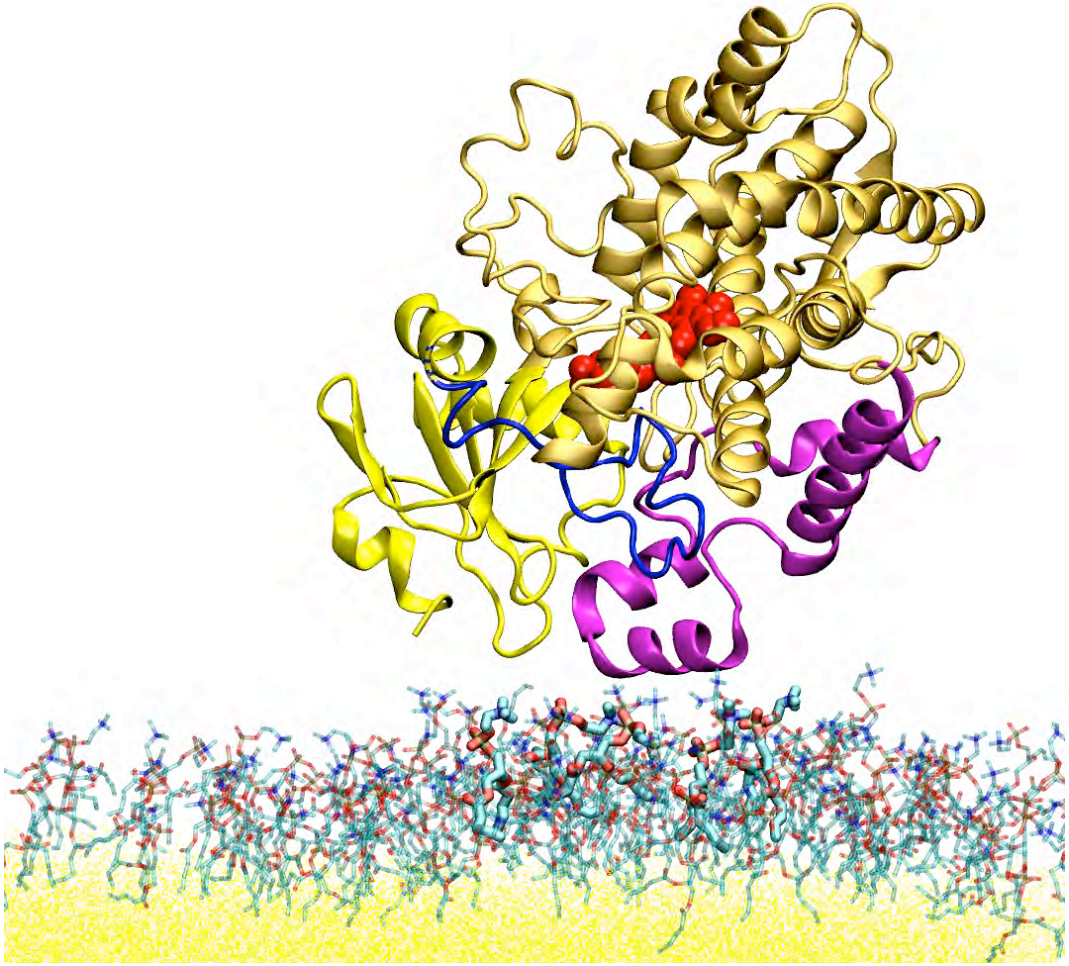


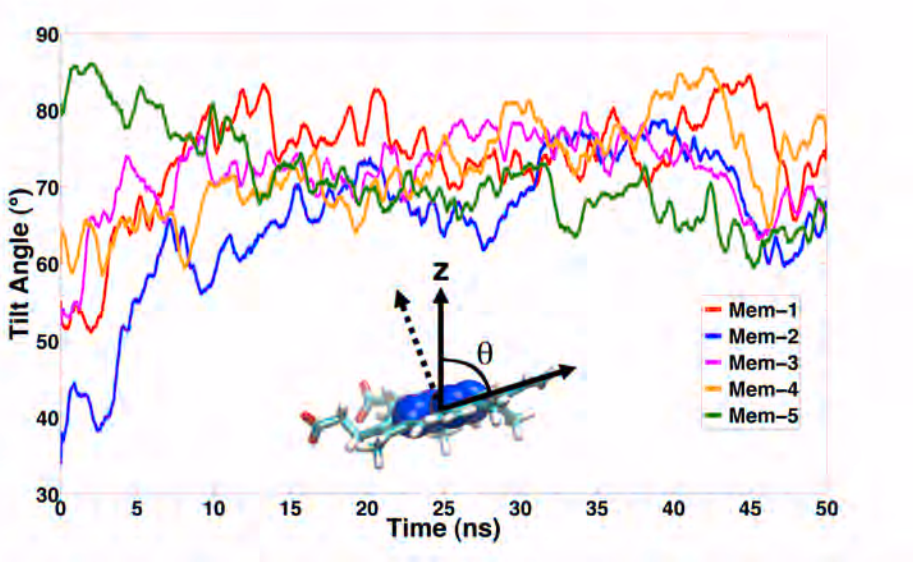


J. Baylon, I. Lenov, S. Sligar and ET, **JACS**, 135: 8542–8551 (2013)

Insertion and Membrane-Induced Conformational Change of Cytochrome P450

Javier Baylon



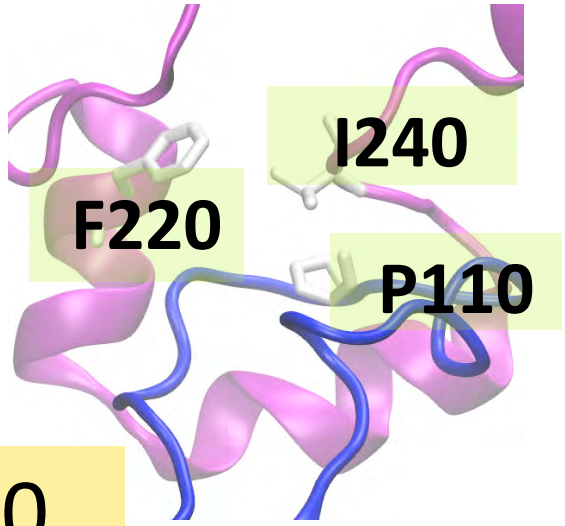


Within 10 degrees of experimental measurement of the tilt angle (S. Sligar)

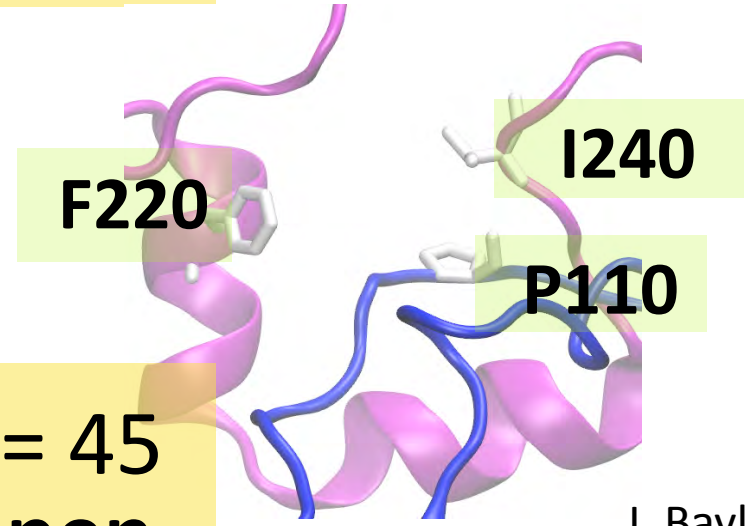
J. Baylon, I. Lenov, S. Sligar and ET, JACS, 135: 8542–8551 (2013)

Insertion and Membrane-Induced Conformational Change of Cytochrome P450

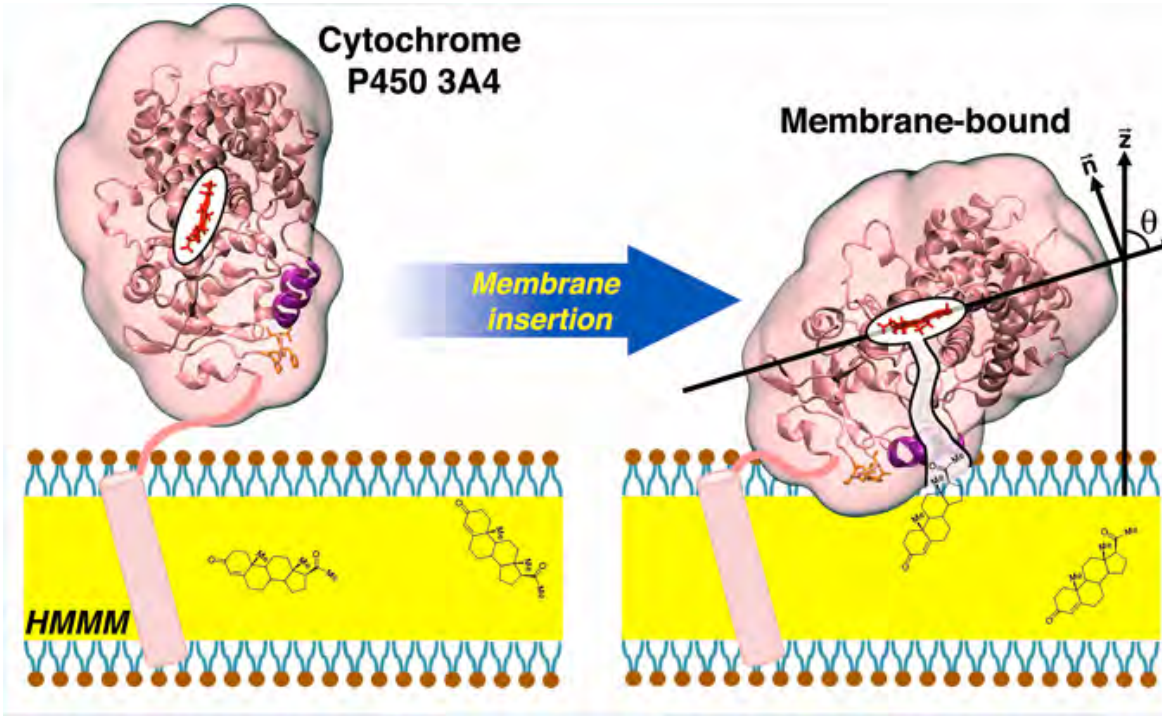
Javier Baylon



$t = 0$
closed



$t = 45$
open



J. Baylon, I. Lenov, S. Sligar and ET, JACS, 135: 8542–8551 (2013)