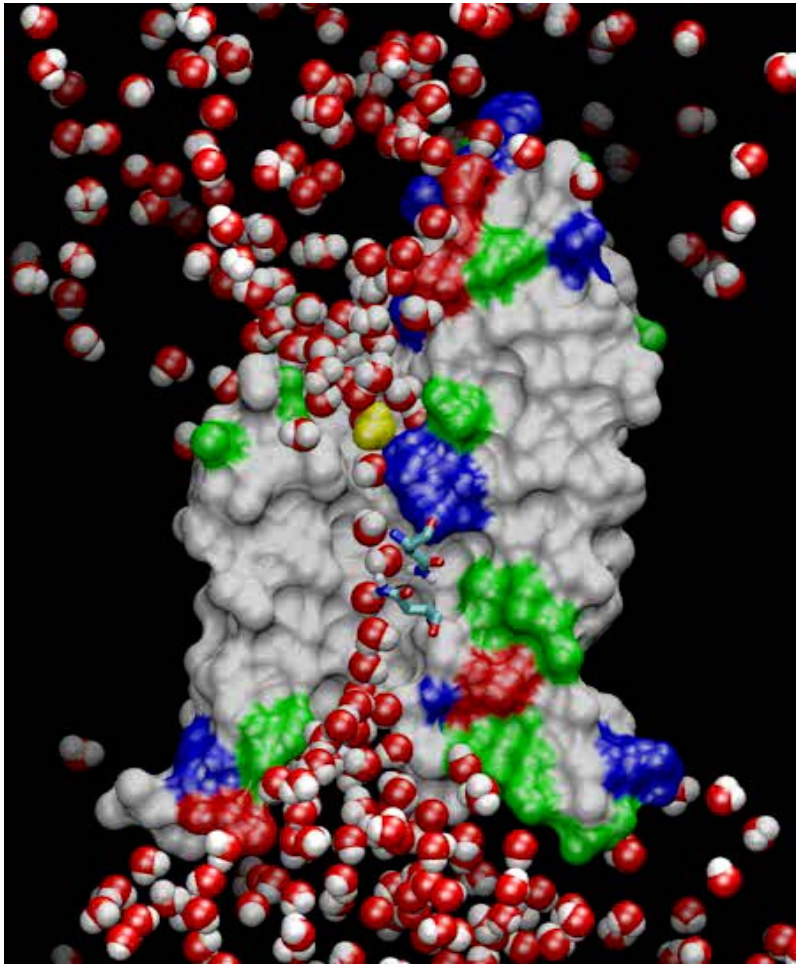
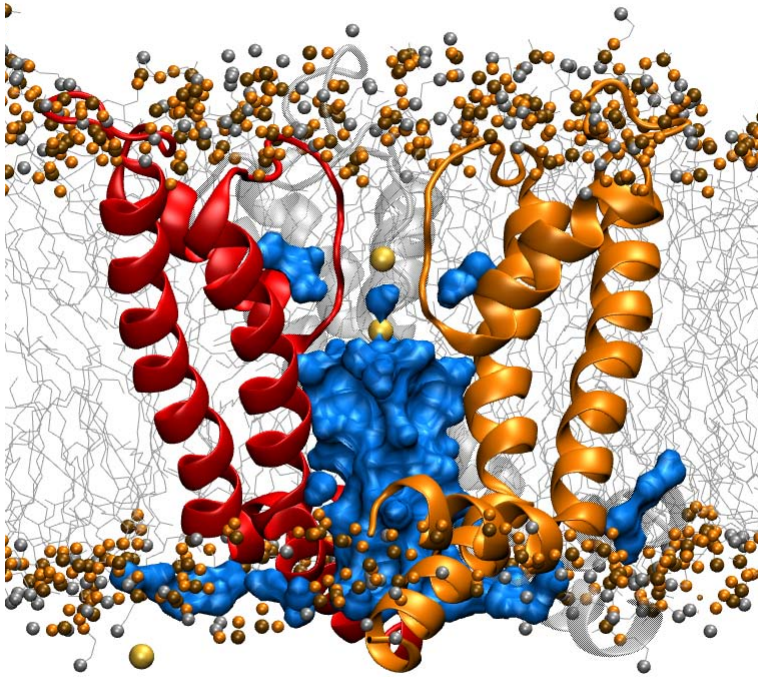
Modeling and Molecular Dynamics of Membrane Proteins



Emad Tajkhorshid

Department of Biochemistry, Center for Biophysics and Computational Biology, and Beckman Institute

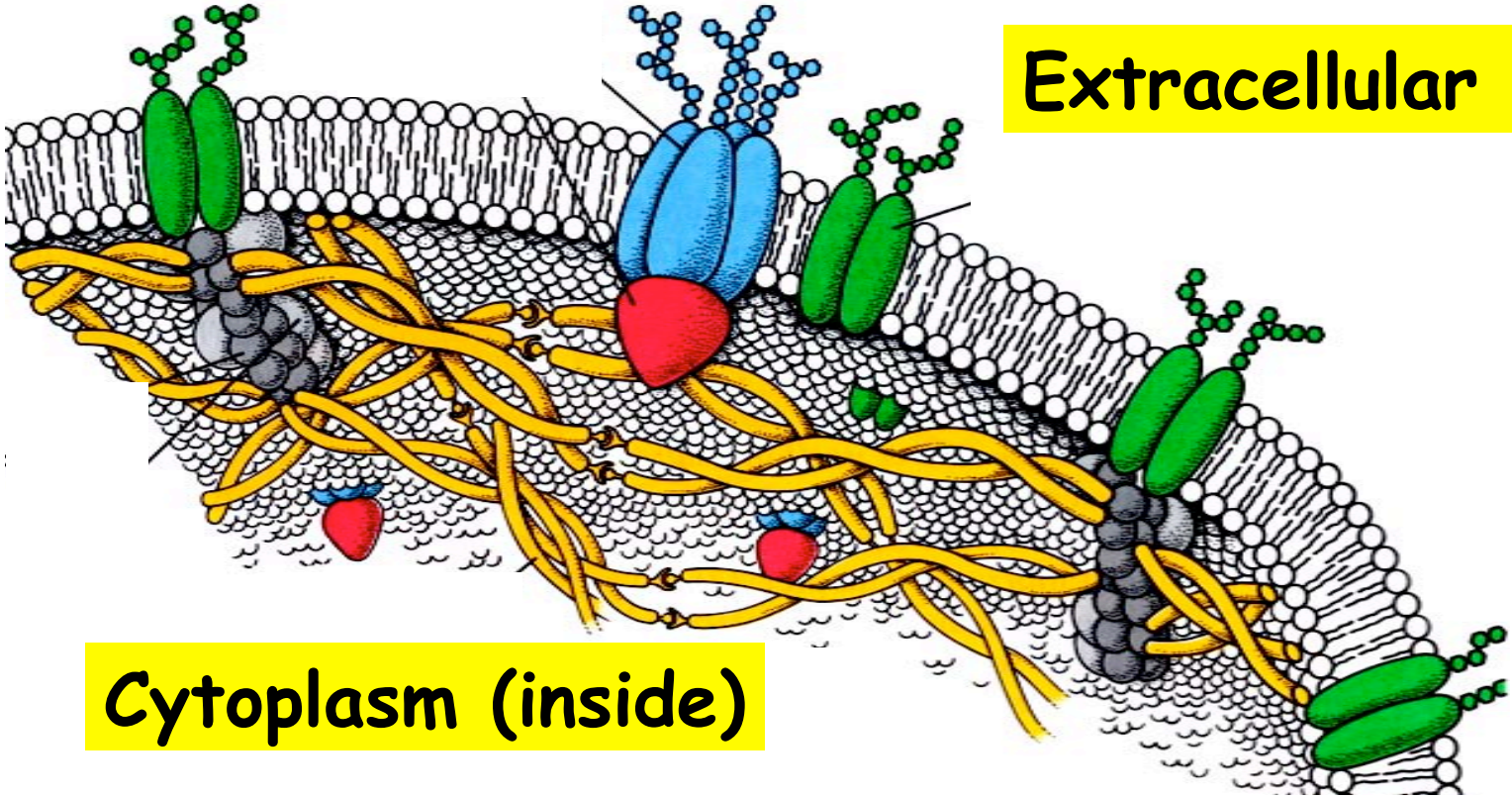
University of Illinois at Urbana-Champagin



Why Do Living Cells Need Membrane Channels (Proteins)?

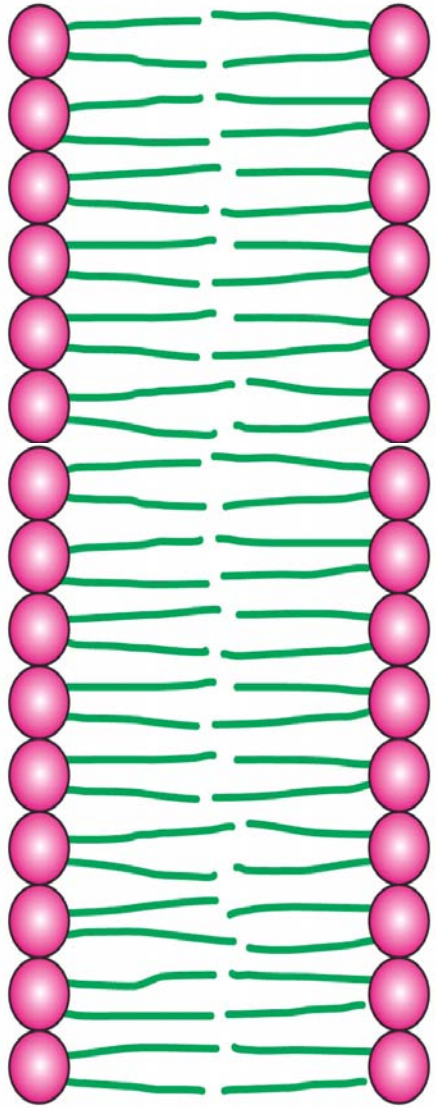
- **Living cells also need to exchange materials and information with the outside world**

… however, in a highly <u>selective</u> manner.

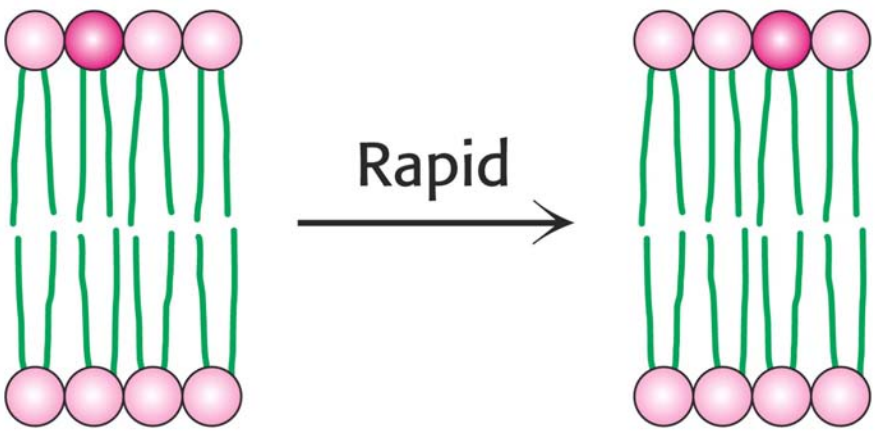


Phospholipid Bilayers Are Excellent Materials For Cell Membranes

- **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



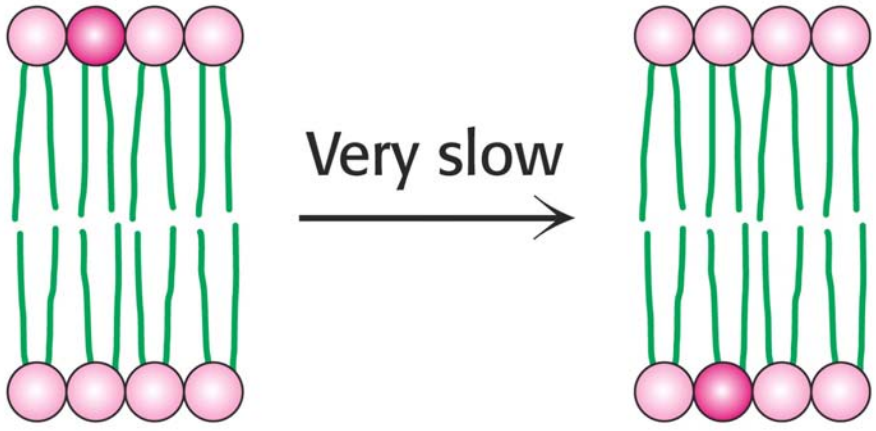
Lateral diffusion

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

(50 Å in ~ 5 x 10^{-6} s)

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

Modeling mixed lipid bilayers!



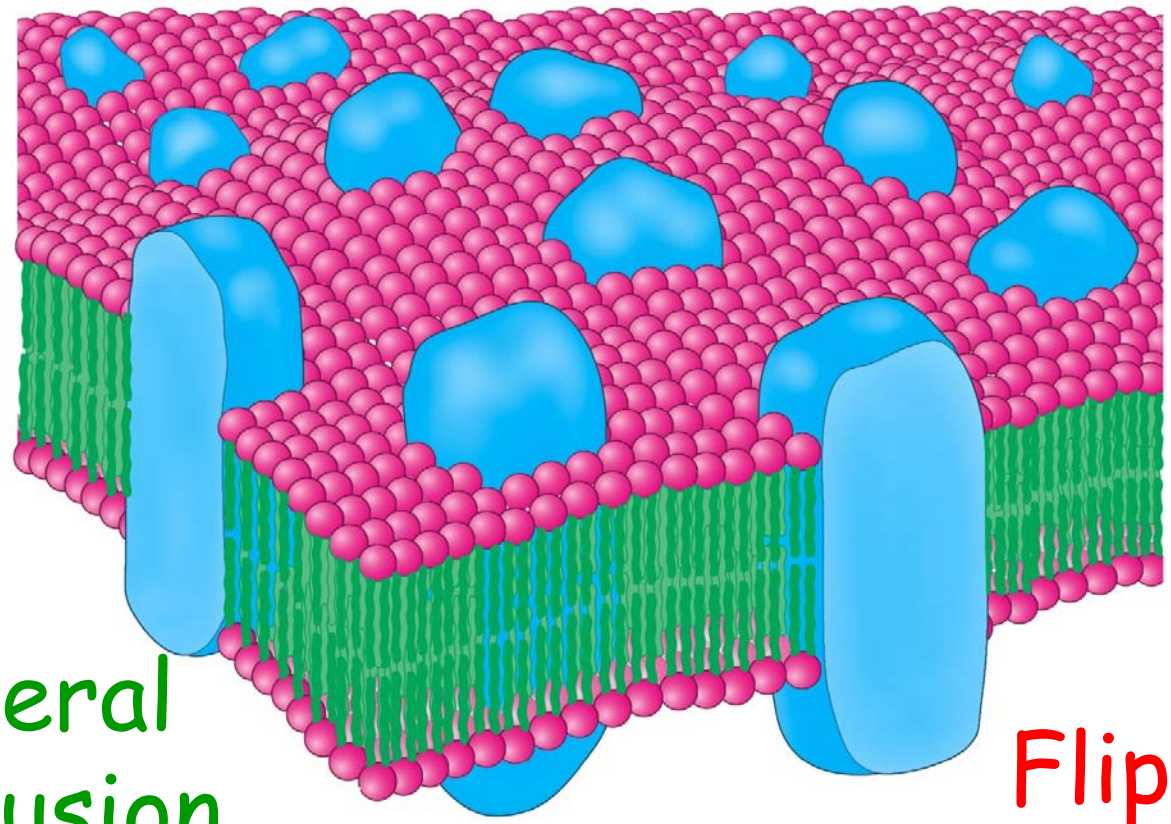
Tranverse diffusion
(flip-flop)

Once in several hours!

(~ 50 Å in ~ 10^4 s)

~9 orders of magnitude slower
ensuring bilayer asymmetry

Fluid Mosaic Model of Membrane



Lateral Diffusion Allowed

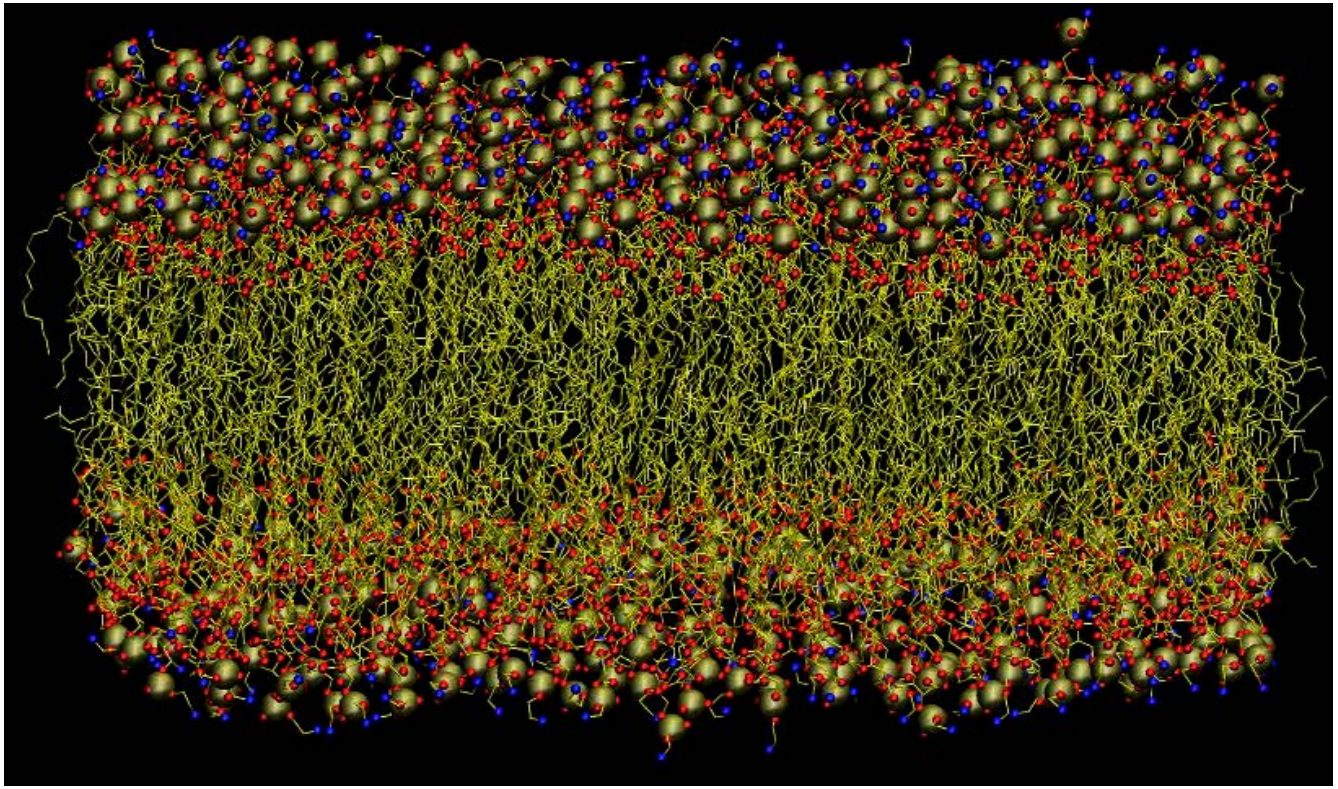
Flip-flop Forbidden

Ensuring the conservation of membrane asymmetric structure

Technical difficulties in Simulations of Biological Membranes

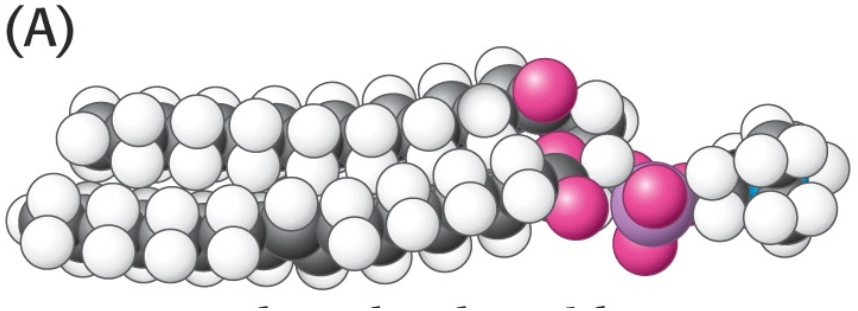
- Time scale
- Heterogeneity of biological membranes ☹

60 x 60 Å
Pure POPE
5 ns
~100,000
atoms



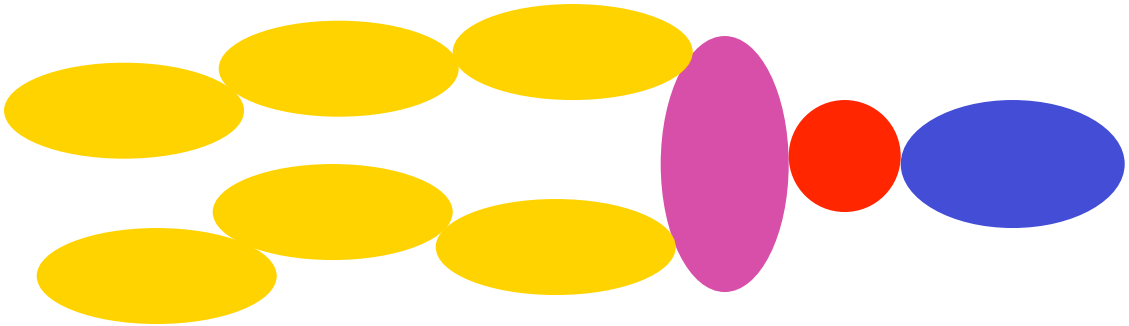
Coarse-grained modeling of lipids

150 particles

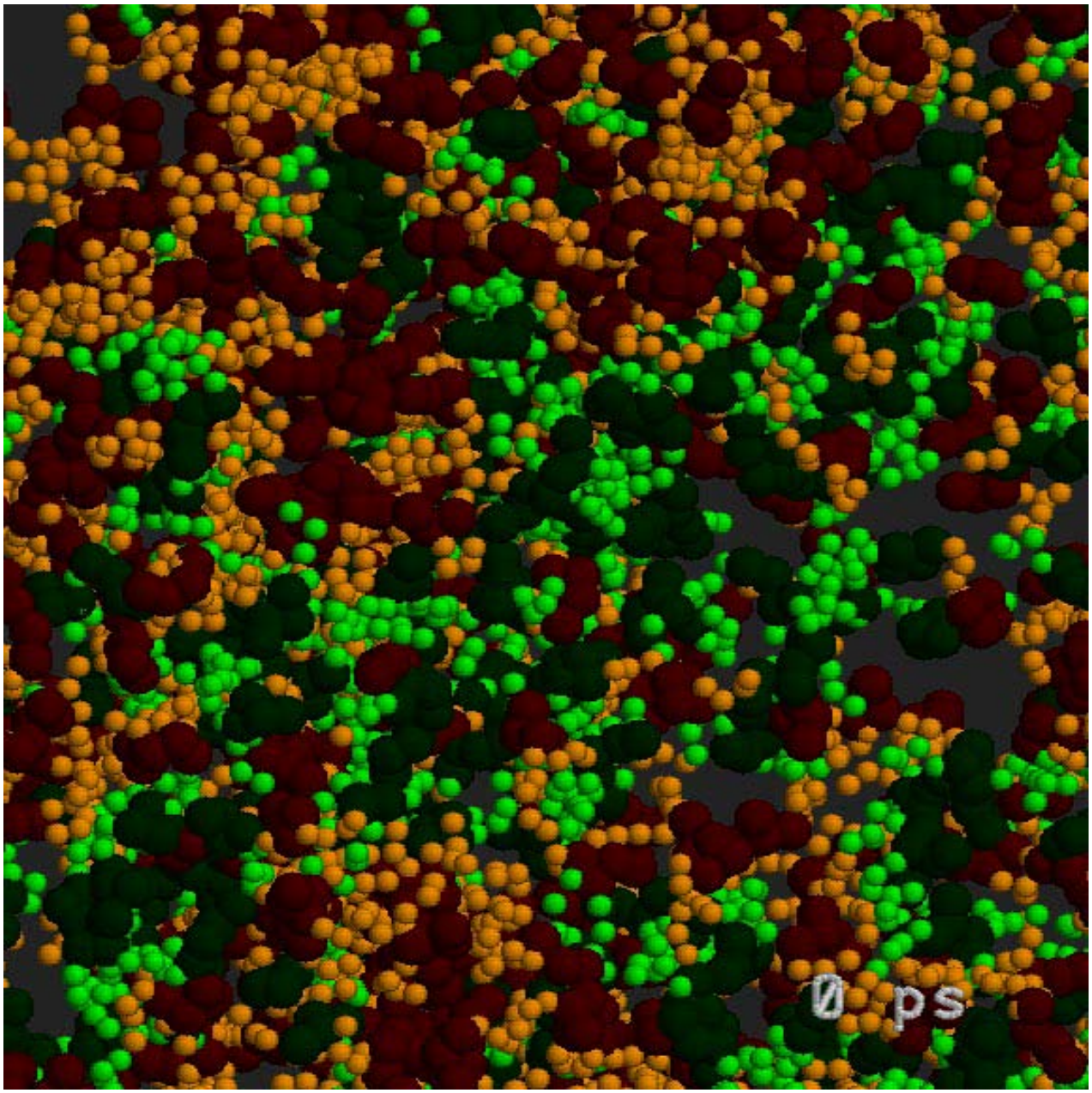


▼

9 particles!



Also, increasing the time step by orders of magnitude.



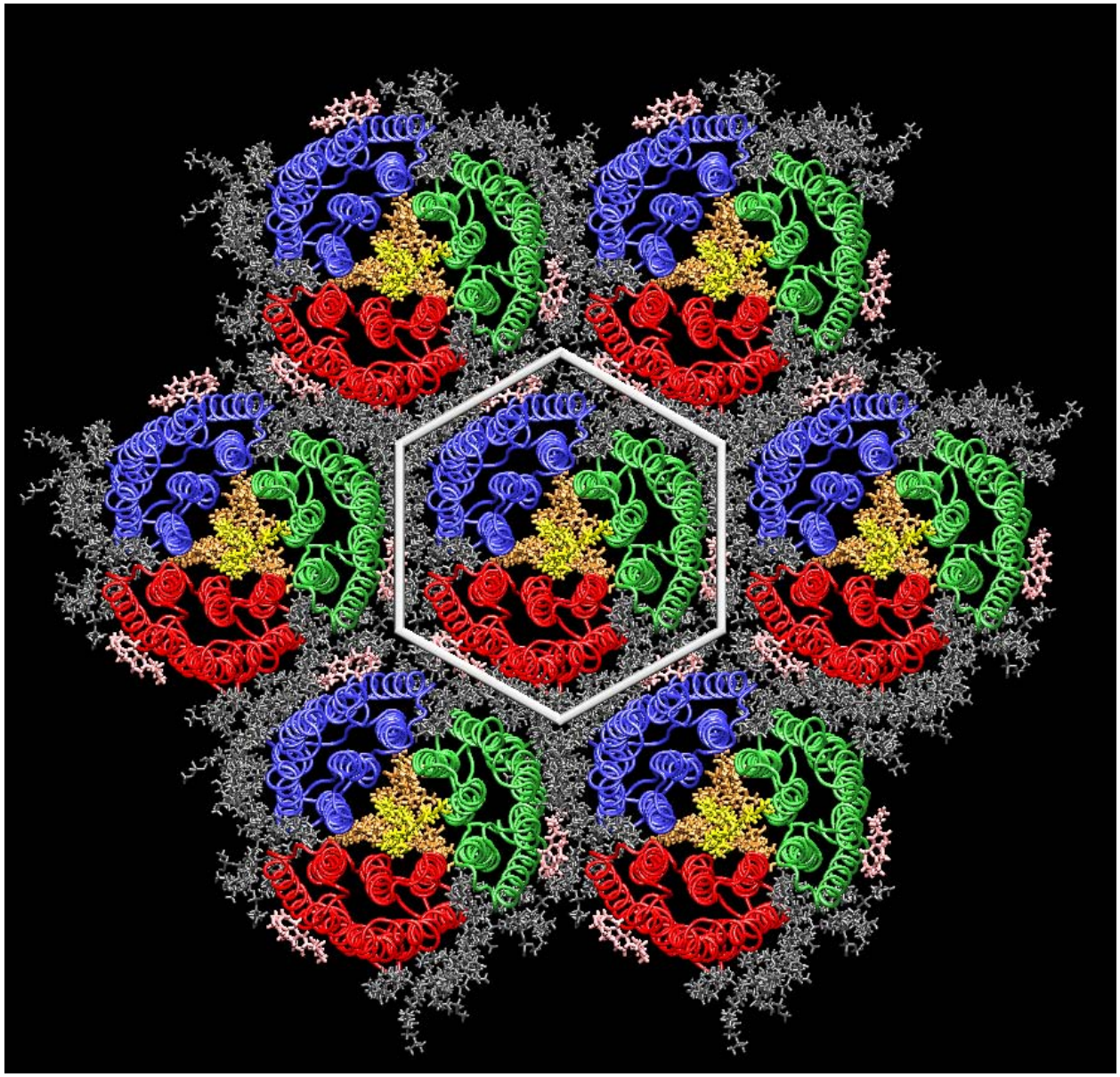
by: J. Siewert-Jan Marrink and Alan E. Mark, University of Groningen, The Netherlands

Protein/Lipid ratio

- Pure lipid: insulation (neuronal cells)
- Other membranes: on average 50%
- Energy transduction membranes (75%)
  Membranes of mitocondria and chloroplast
  Purple membrane of halobacteria

- Different functions = different protein composition

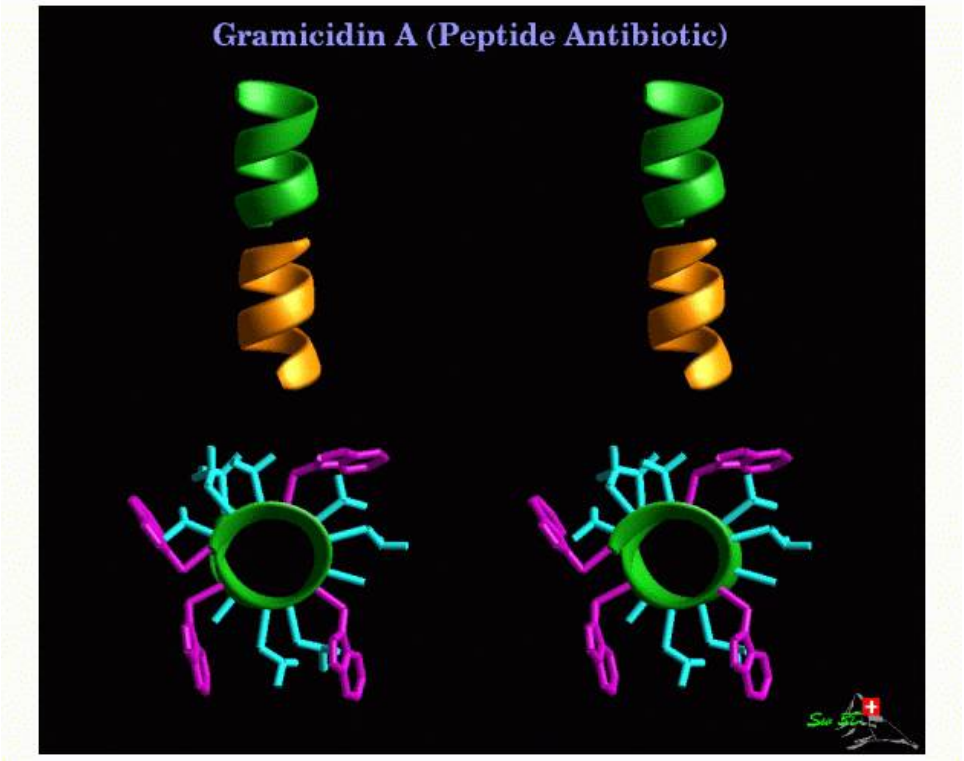
Protein / Lipid Composition

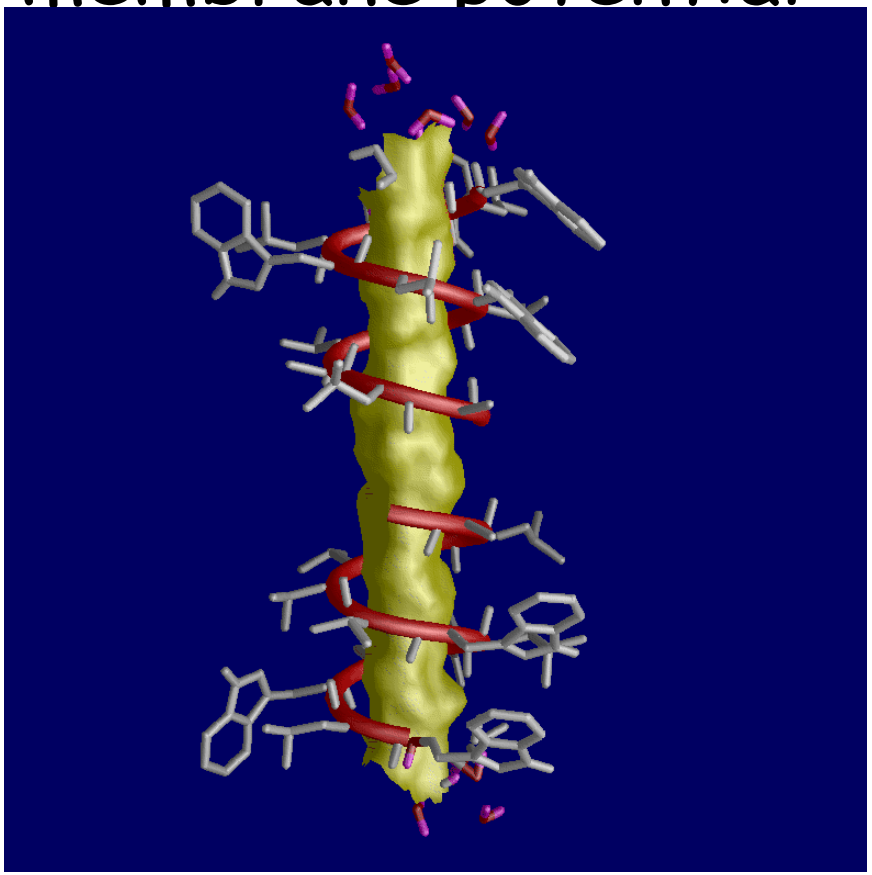


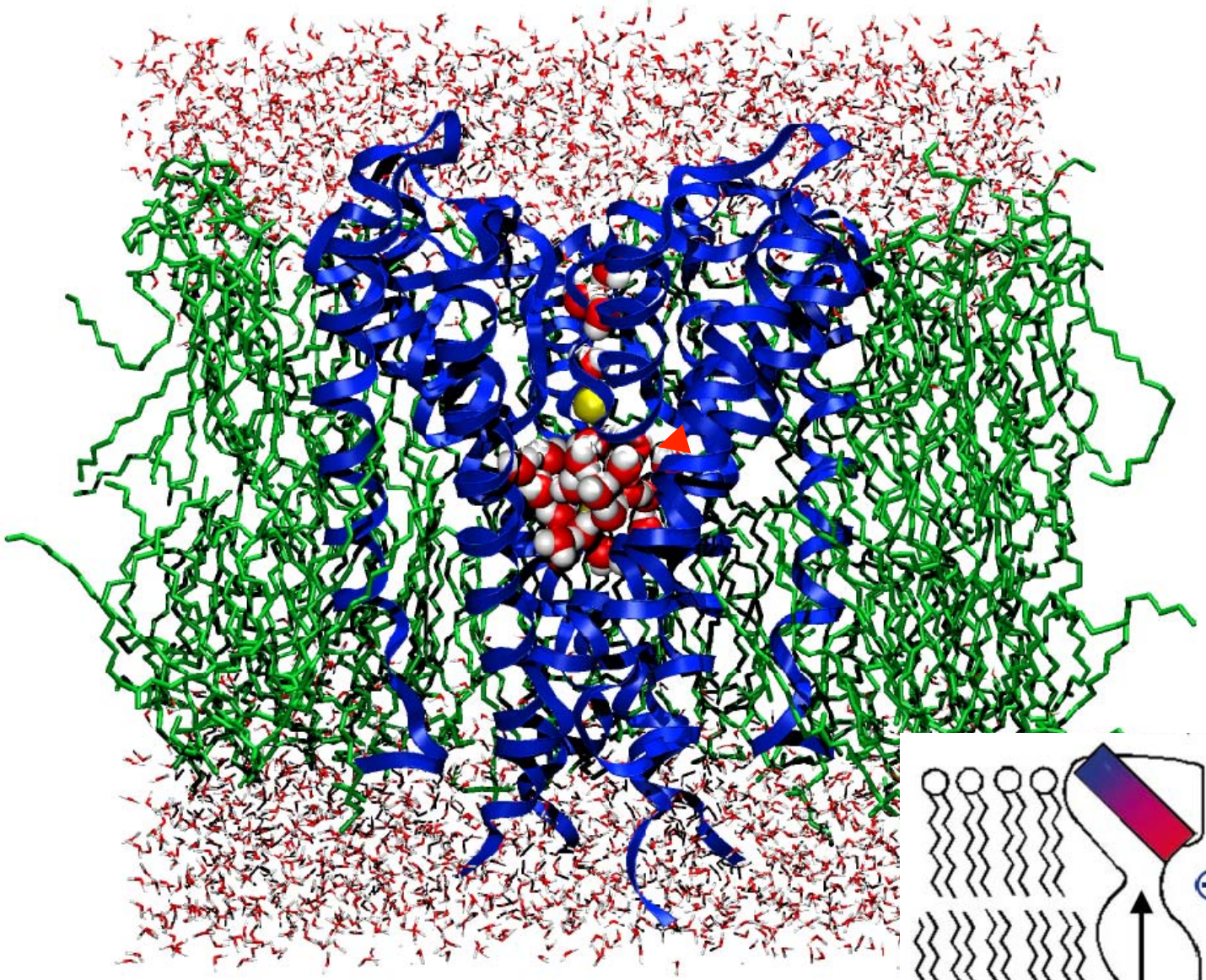
The purple membrane of halobacteria

Gramicidin A

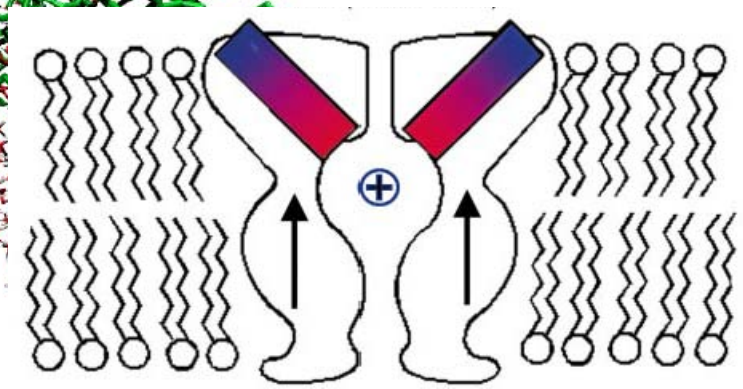
Might be very sensitive to the lipid head group electrostatic and membrane potential







Central cavity

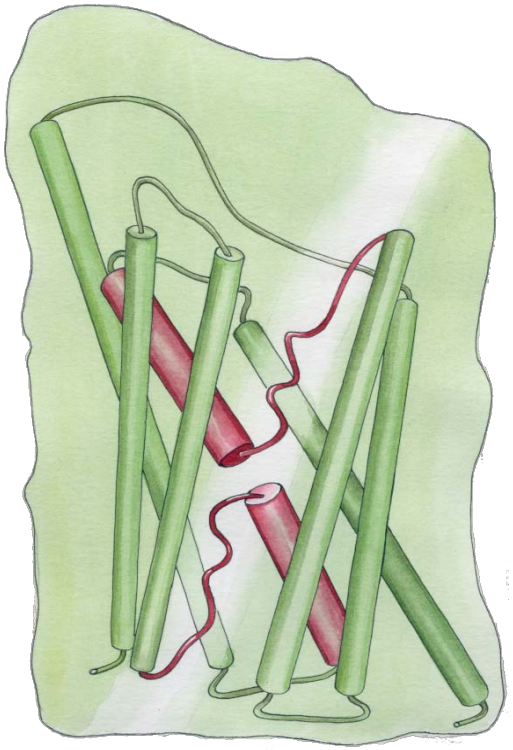


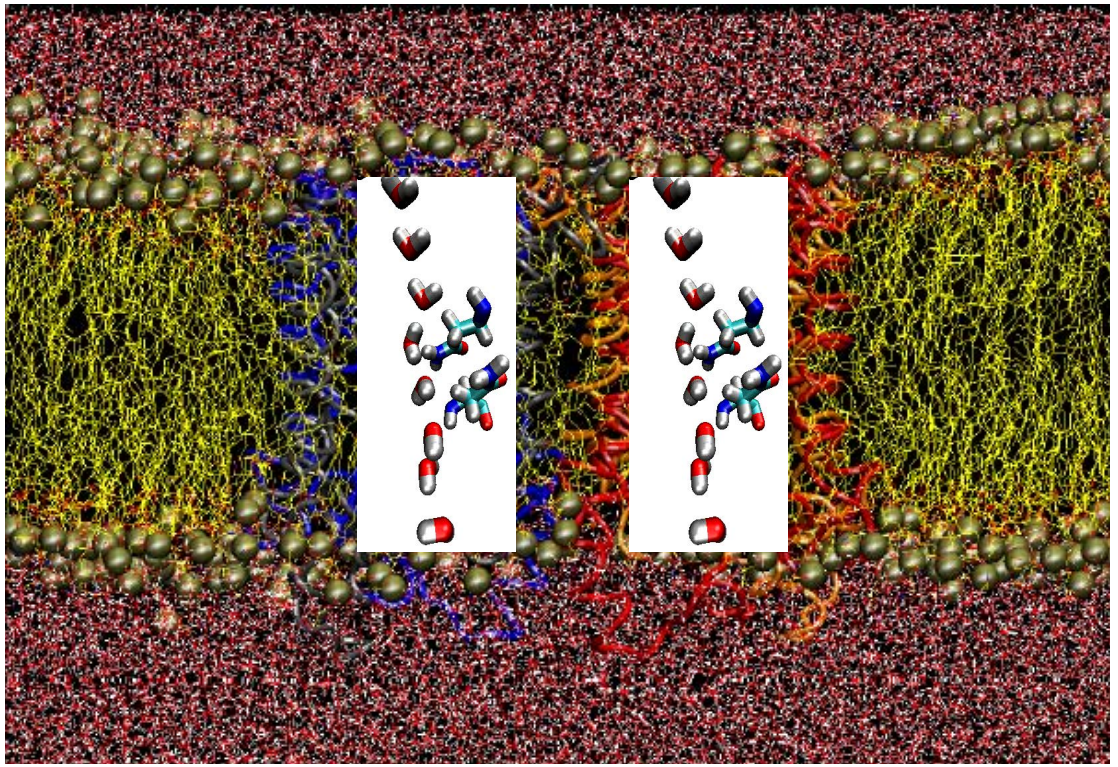
Analysis of Molecular Dynamics Simulations of Biomolecules

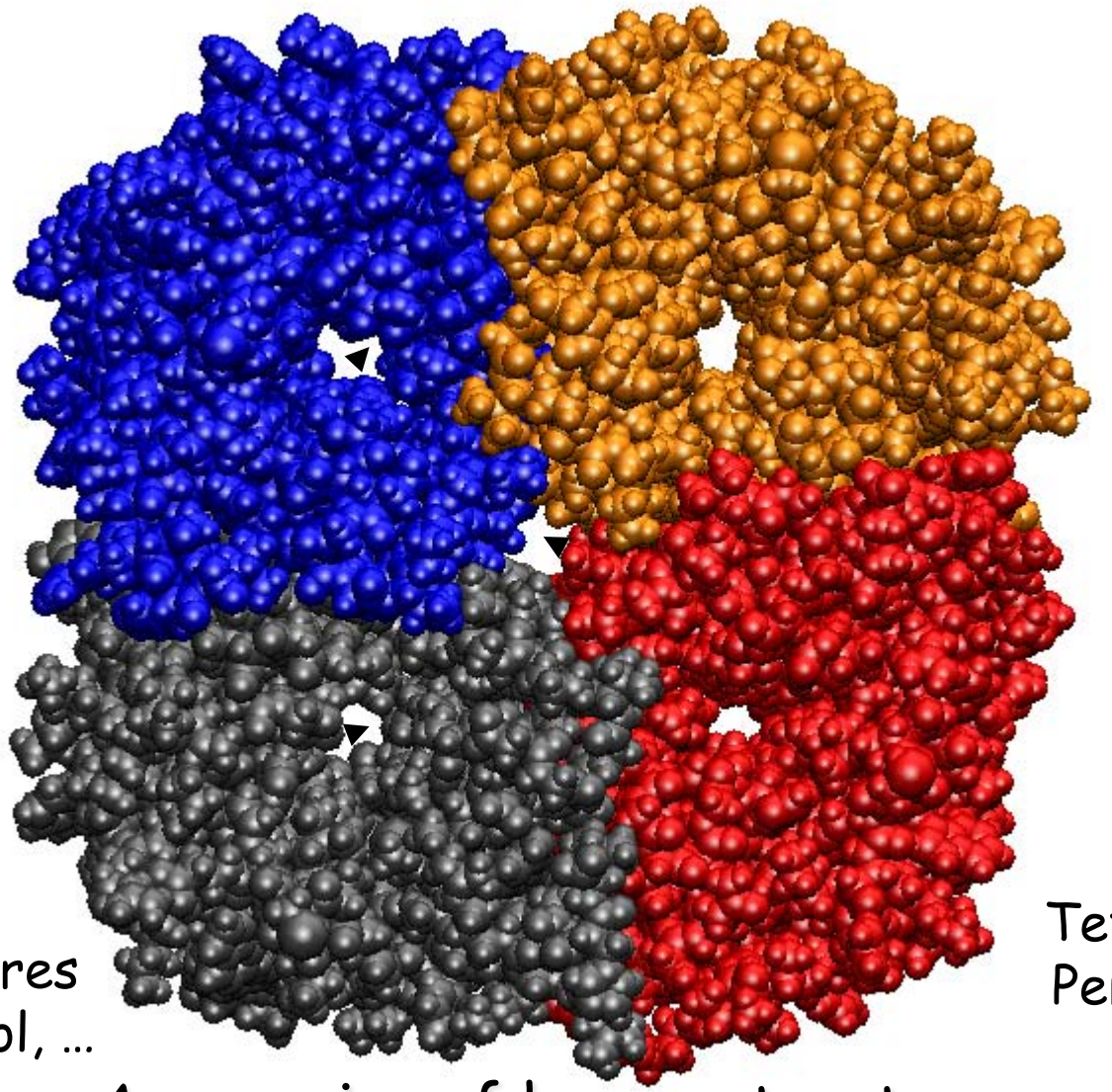
- A very complicated arrangement of hundreds of groups interacting with each other
- Where to start to look at?
- What to analyze?
- How much can we learn from simulations?

It is very important to get acquainted with your system

Aquaporins
Membrane water channels







Monomeric pores
Water, glycerol, …

Tetrameric pore
Perhaps ions???

Aquaporins of known structure:

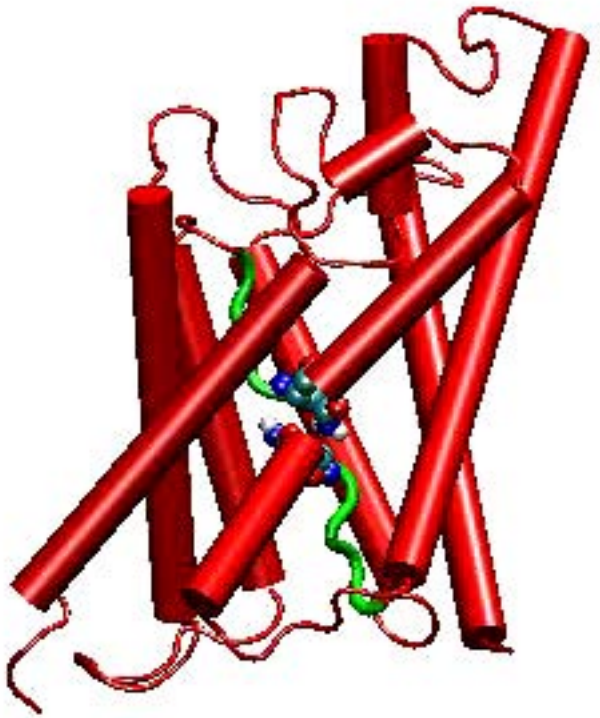
GlpF – E. coli glycerol channel (aquaglycerolporin)

AQP1 – Mammalian aquaporin-1 (pure water channel)

AqpZ and AQP0 (2004)

Functionally Important Features

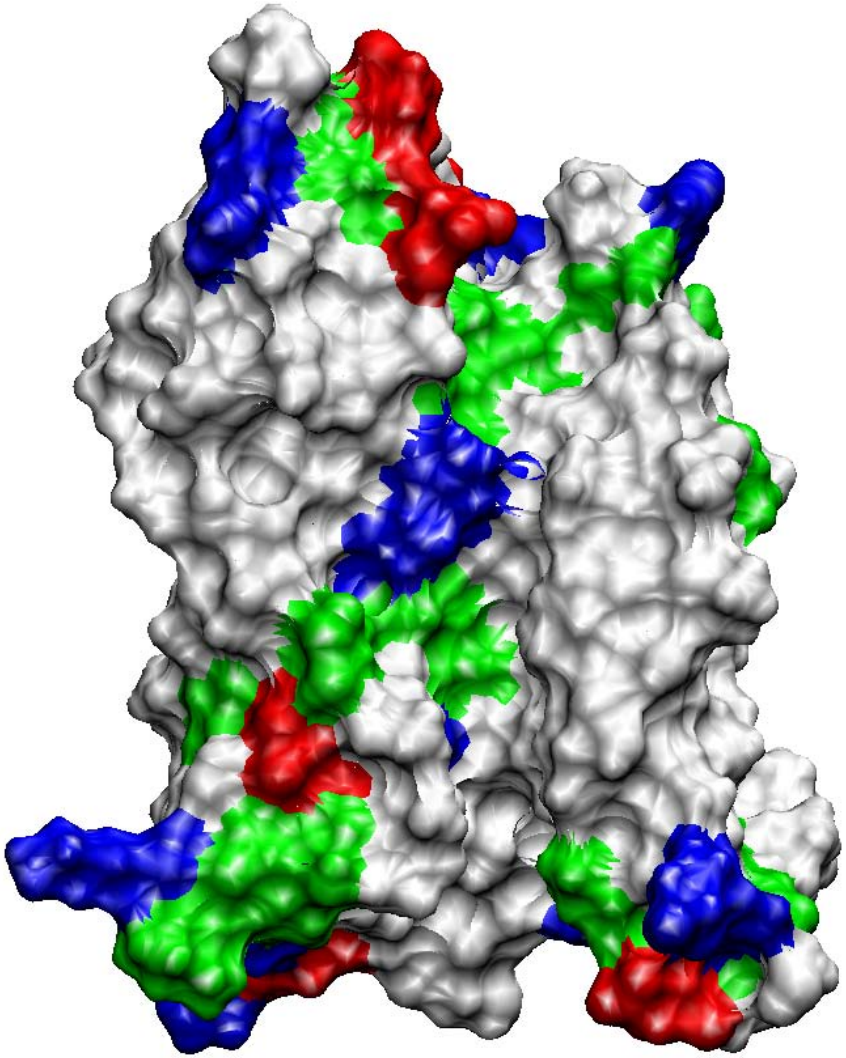
- Tetrameric architecture
- Amphipatic channel interior
- Water and glycerol transport
- Protons, and other ions are excluded
- Conserved asparagine-proline-alanine residues; NPA motif
- Characteristic half-membrane spanning structure



~100% conserved -NPA- signature sequence

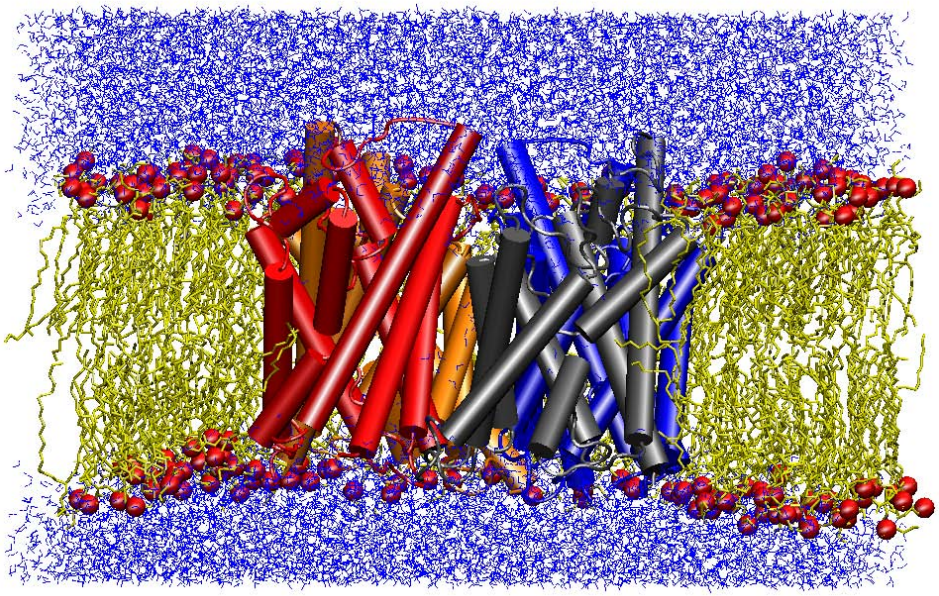
N **E NPA** **E NPAR** C

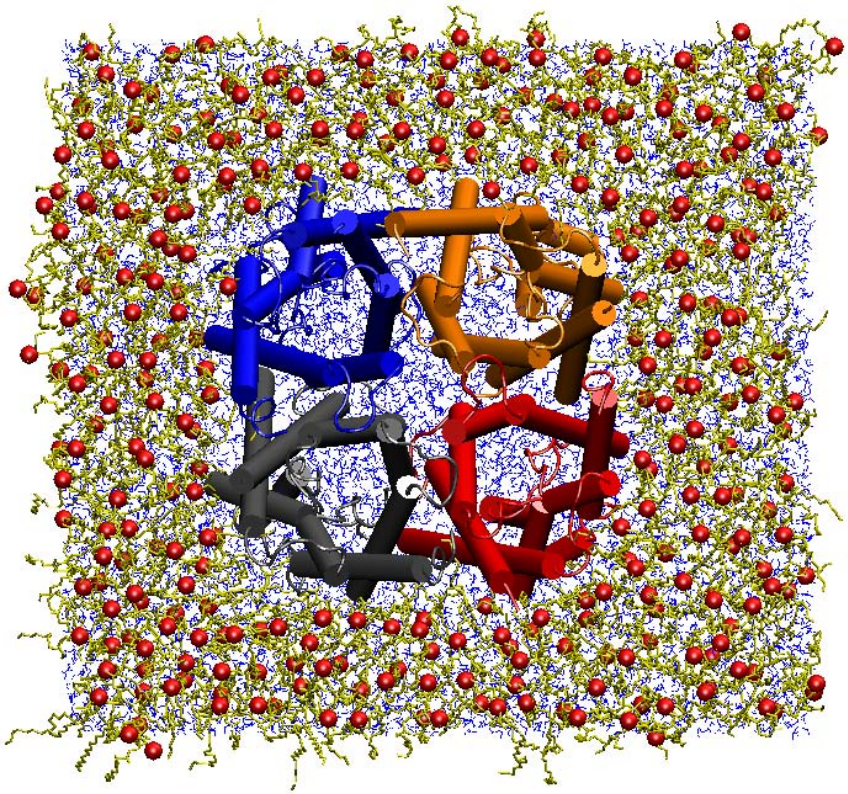
A Semi-hydrophobic channel



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

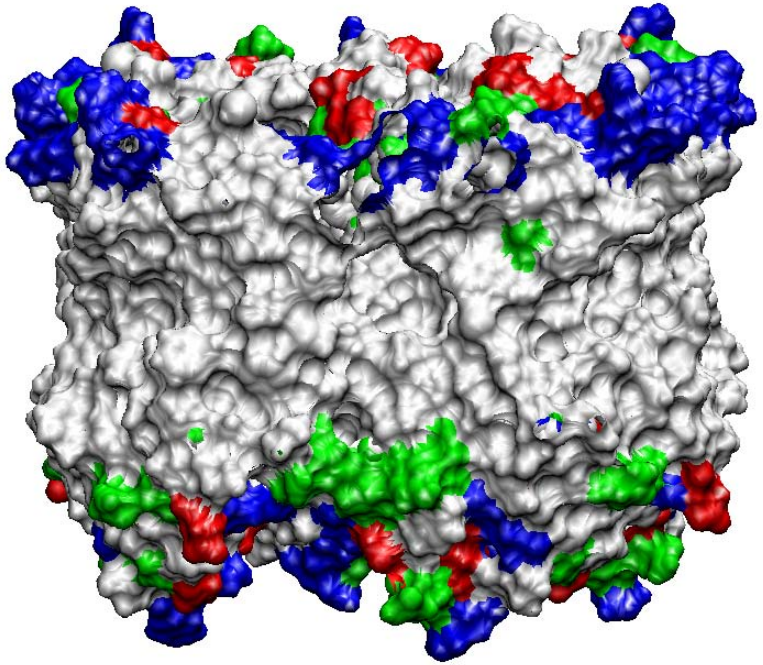
1ns equilibration, 4ns production

10 days /ns – 32-proc Linux cluster

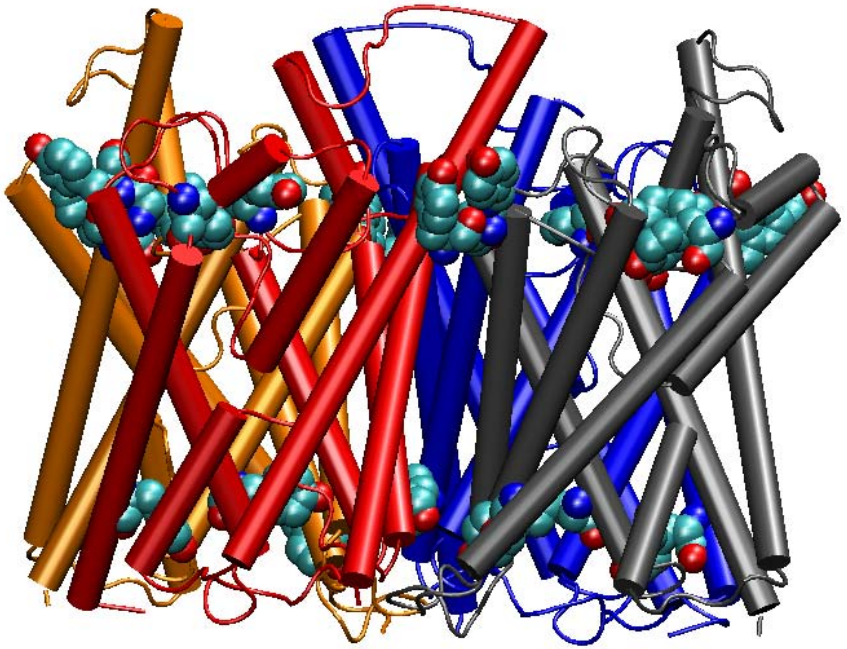
3.5 days/ns - 128 O2000 CPUs

0.35 days/ns – 512 LeMieux CPUs

Protein Embedding in Membrane



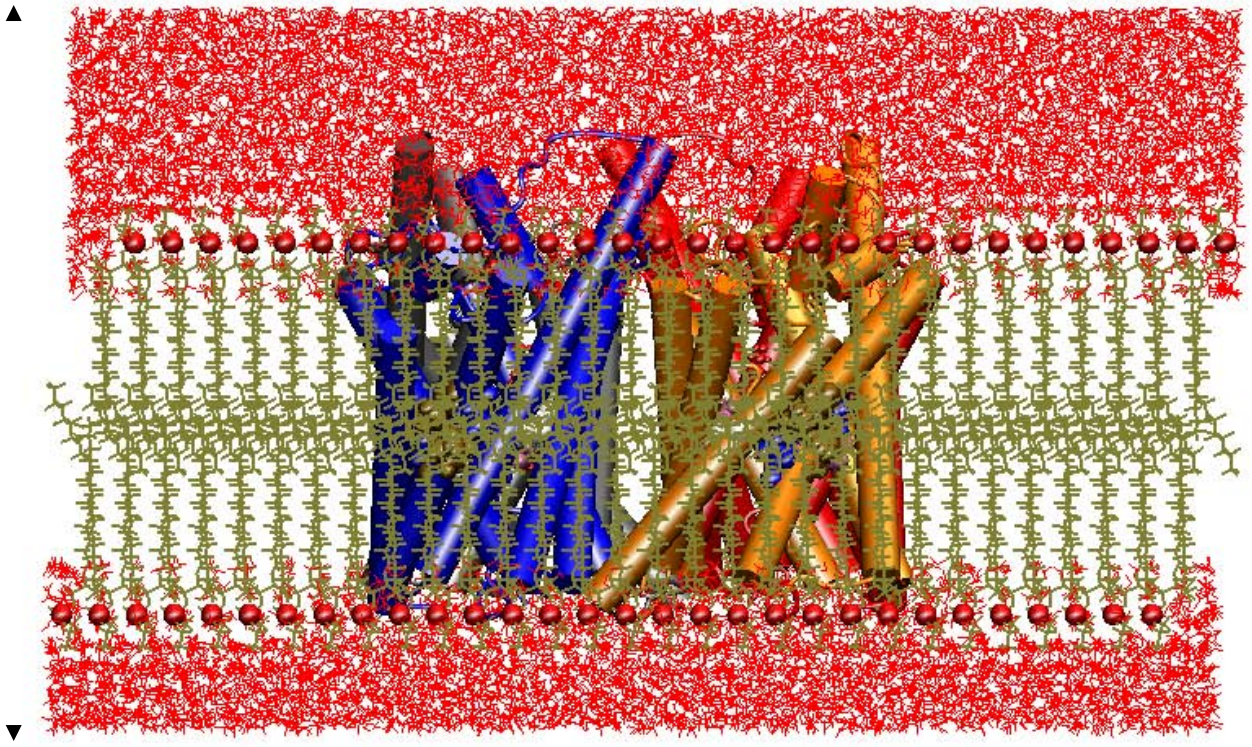
Hydrophobic surface
of the protein

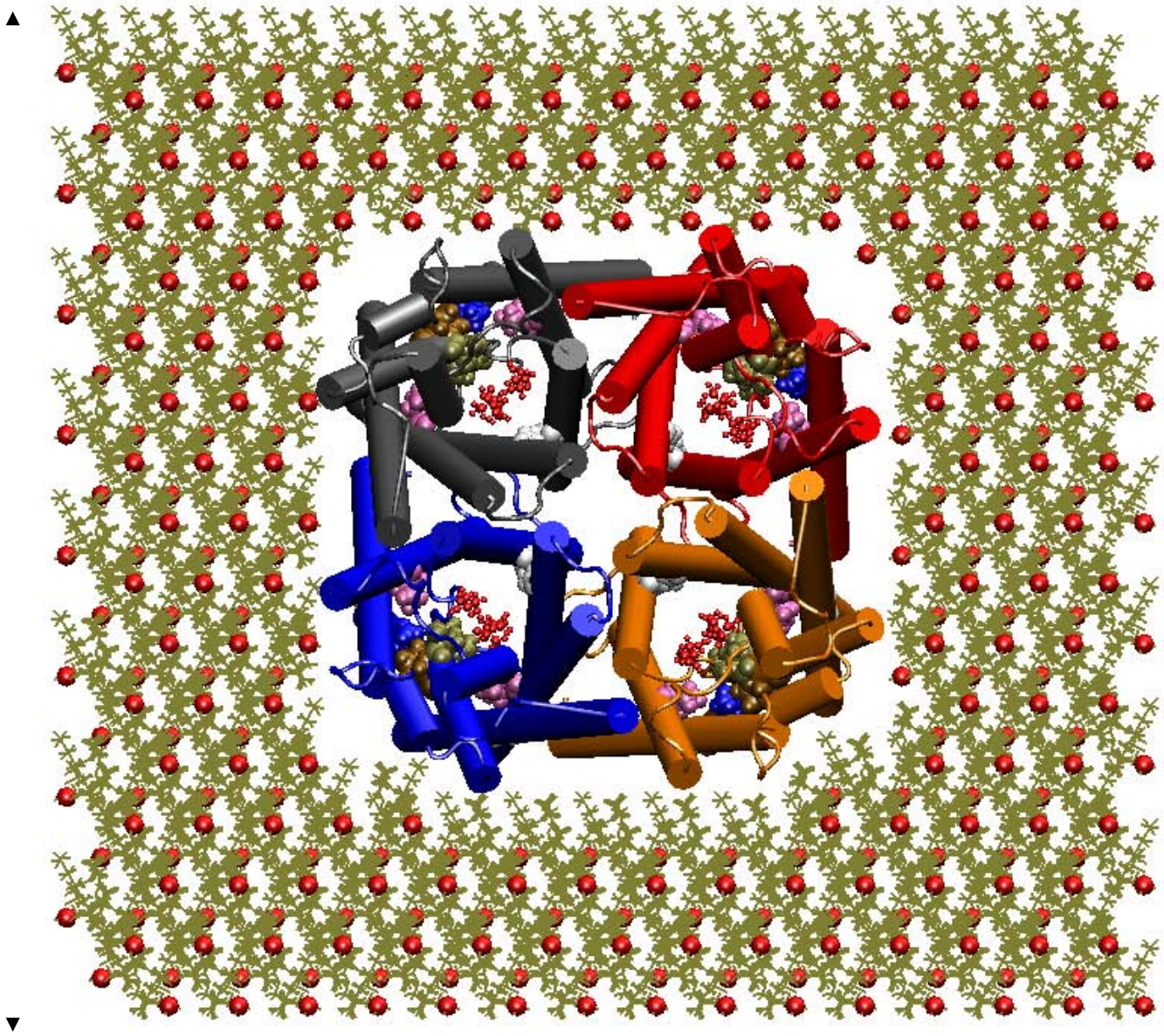


Ring of
Tyr and Trp

Embedding GlpF in Membrane

77 A




112 A
122 A

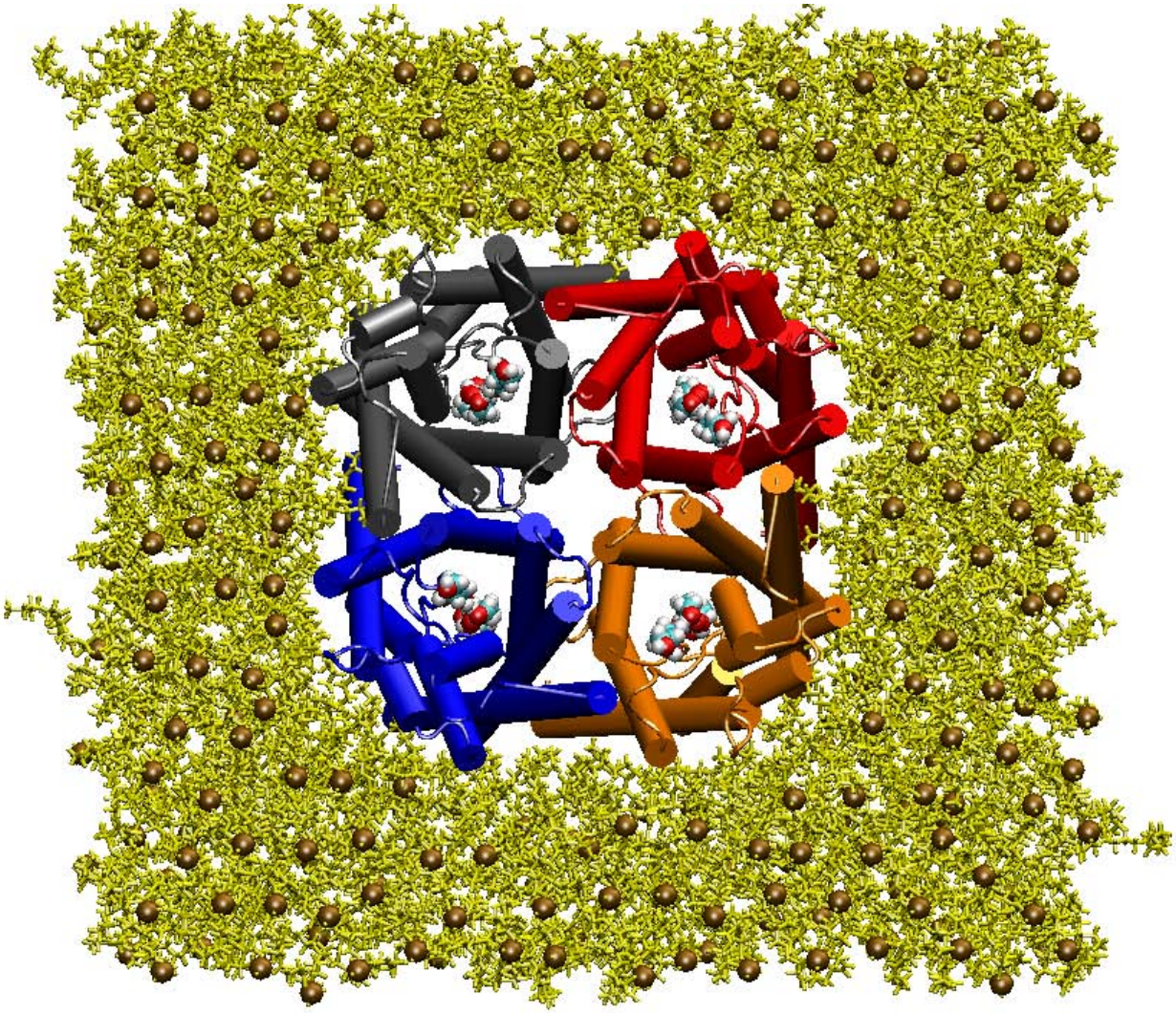
A Recipe for Membrane Protein Simulations

- Align the protein along the *z*-axis (membrane normal): OPM, Orient.
- Decide on the lipid type and generate a large enough patch (MEMBRANE plugin in VMD, other sources). Size, area/lipid, shrinking.
- Overlay the protein with a hydrated lipid bilayer. Adjust the depth/ height to maximize hydrophobic overlap and matching of aromatic side chains (**Trp/Tyr**) with the interfacial region
- Remove lipids/water that overlap with the protein. Better to keep as many lipids as you can, so try to remove clashes if they are not too many by playing with the lipids. Add more water and ions to the two sides of the membrane (SOLVATE / AUTOIONIZE in VMD)
- **Constrain** (not **FIX**) the protein (we are still modeling, let's preserve the crystal structure; fix the lipid head groups and water/ion and **minimize/simulate** the lipid tails using a short simulation.

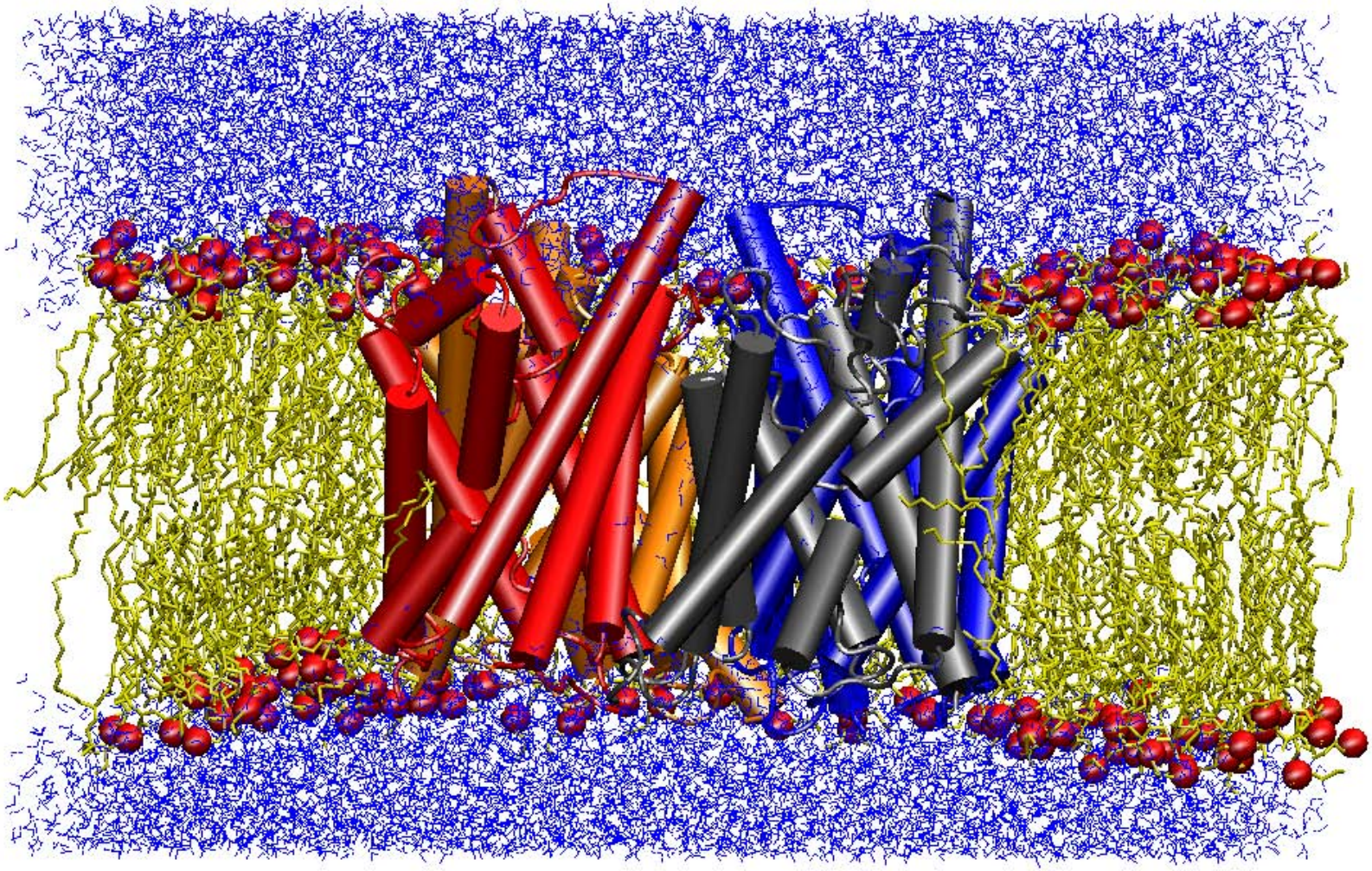
A Recipe for Membrane Protein Simulations

- Continue to constrain the protein (heavy atoms), but release everything else; minimize/simulate using a short "constant-pressure" MD (NPT) to "pack" lipids and water against the protein and fill the gaps introduced after removal of protein-overlapping lipids.
- Watch water molecules; They normally stay out of the hydrophobic cleft. If necessary apply constraints to prevent them from penetrating into the open cleft between the lipids and the protein.
- Monitor the volume of your simulation box until the steep phase of the volume change is complete (.xst and .xsc files). Do not run the system for too long during this phase (over-shrinking; sometimes difficult to judge).
- Now release the protein, minimize the whole system, and start another short NPT simulation of the whole system.
- Switch to an NP_nAT or an NVT simulation, when the system reaches a **stable** volume. Using the new CHARMM force field, you can stay with NPT.

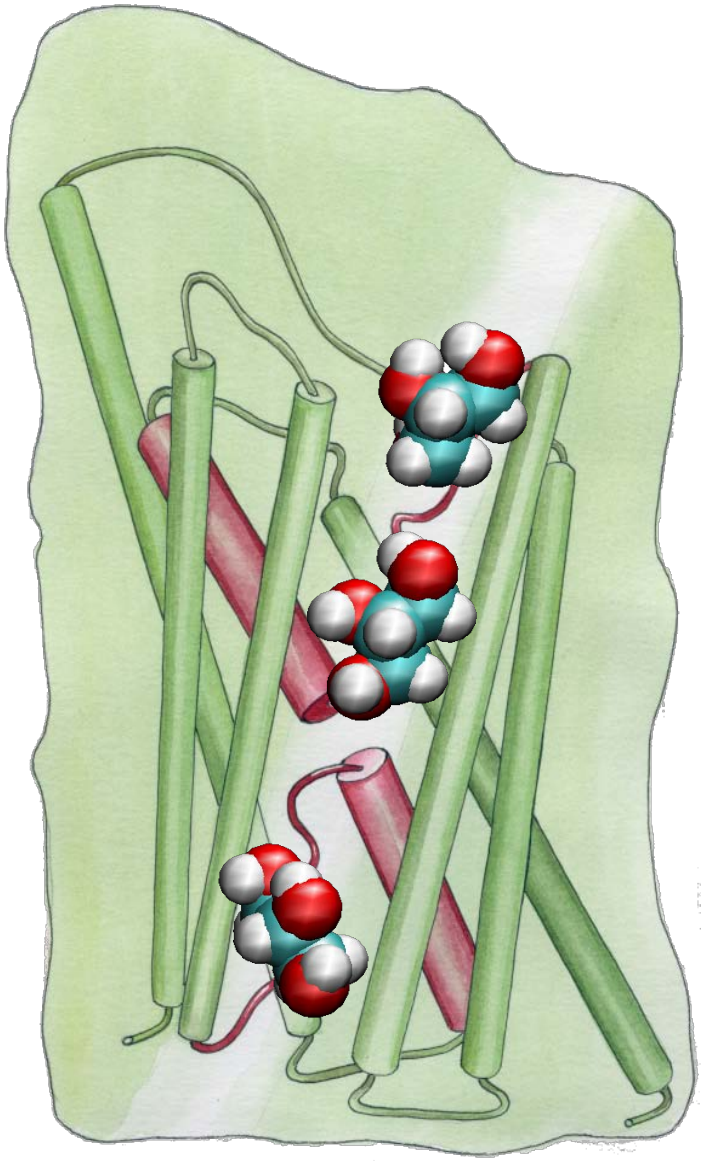
Lipid-Protein Packing During the Initial NpT Simulation

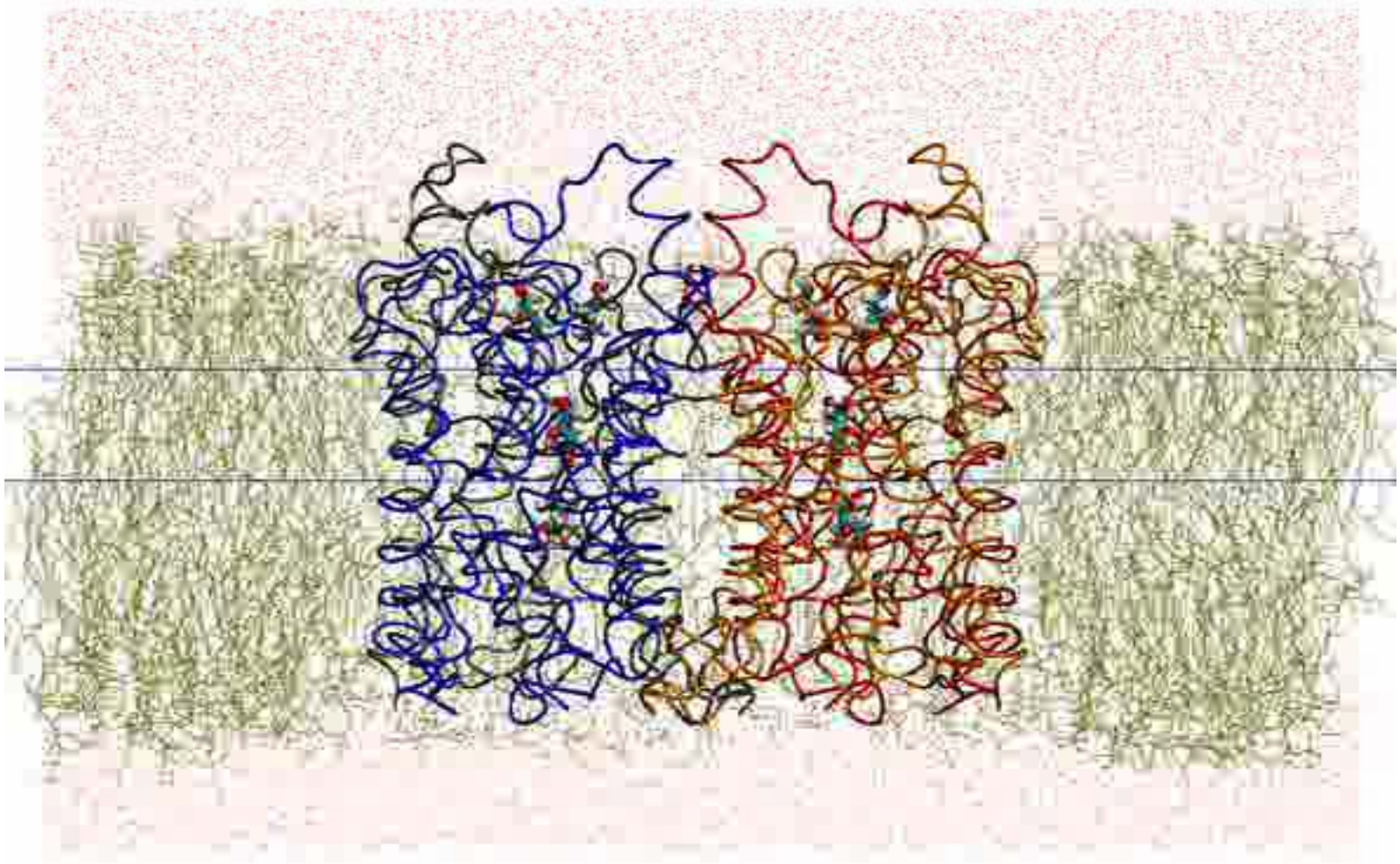


Adjustment of Membrane Thickness to the Protein Hydrophobic Surface

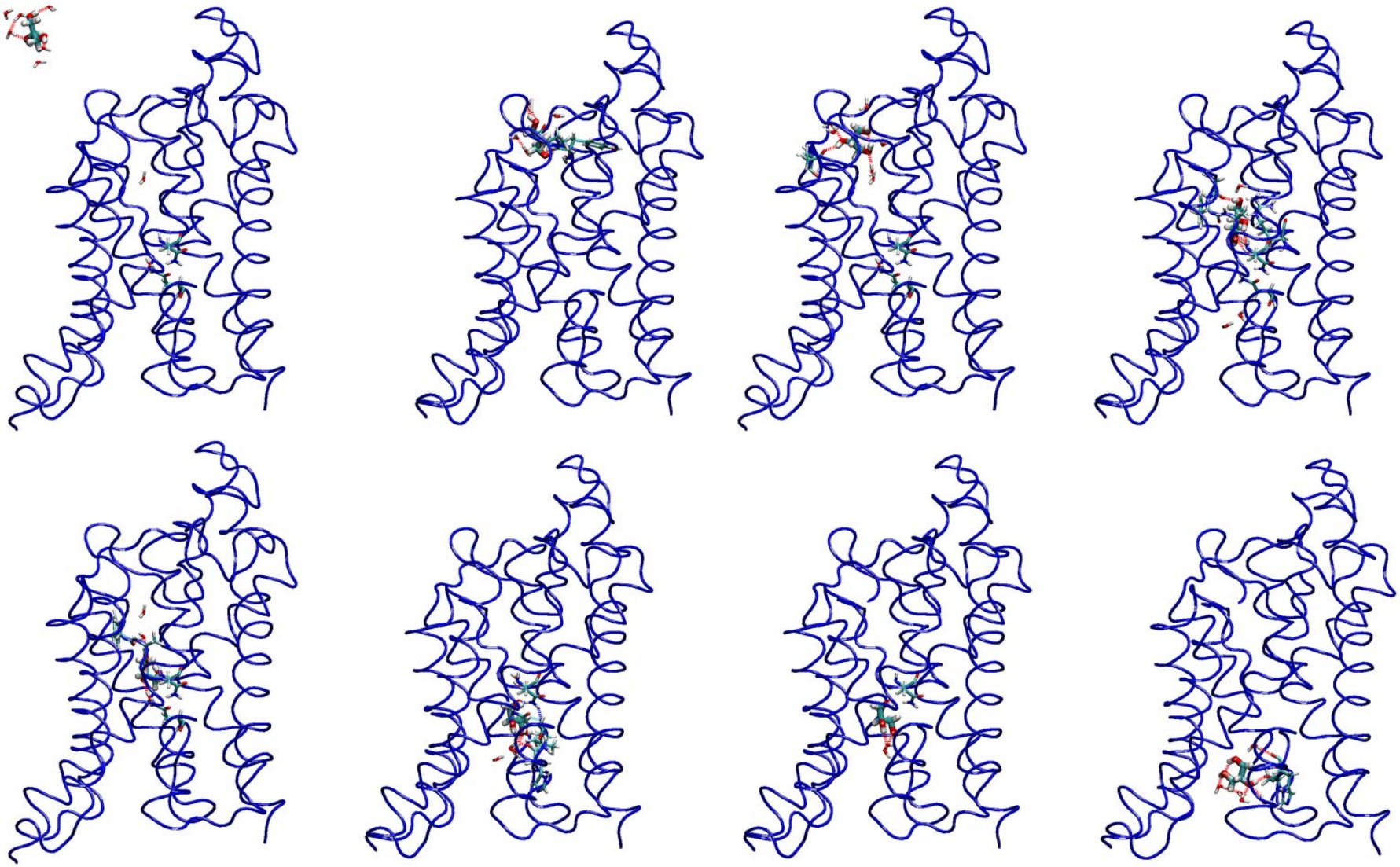


Glycerol-Saturated GlpF



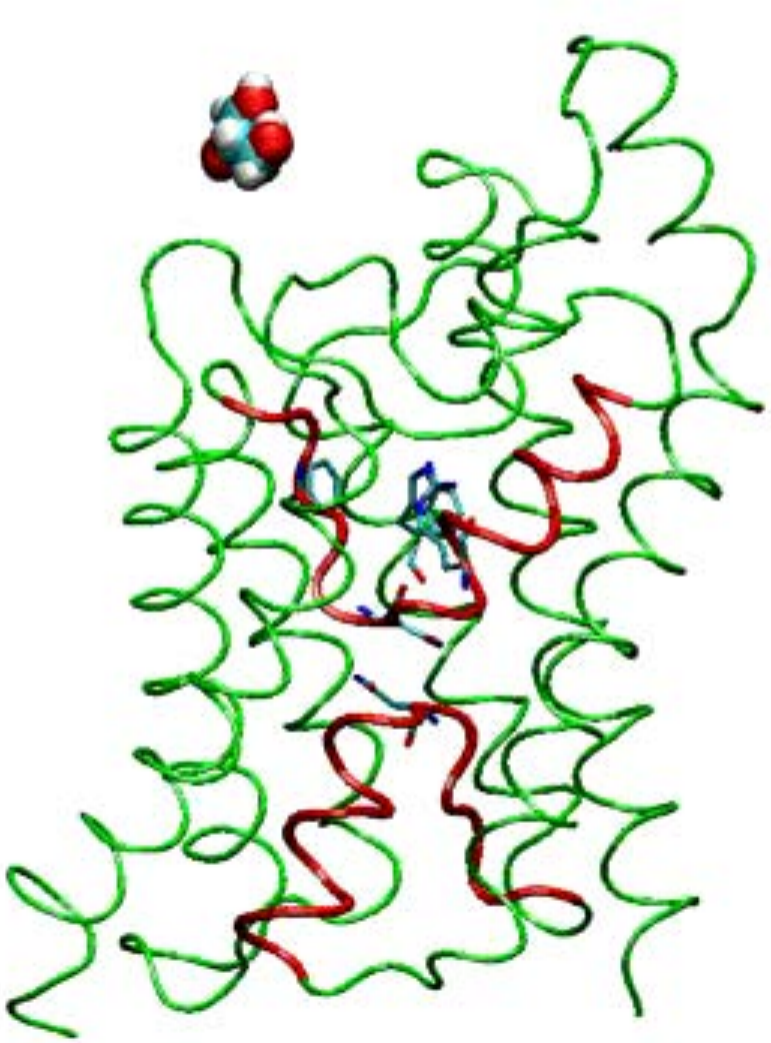


Description of full conduction pathway



Complete description of the conduction pathway

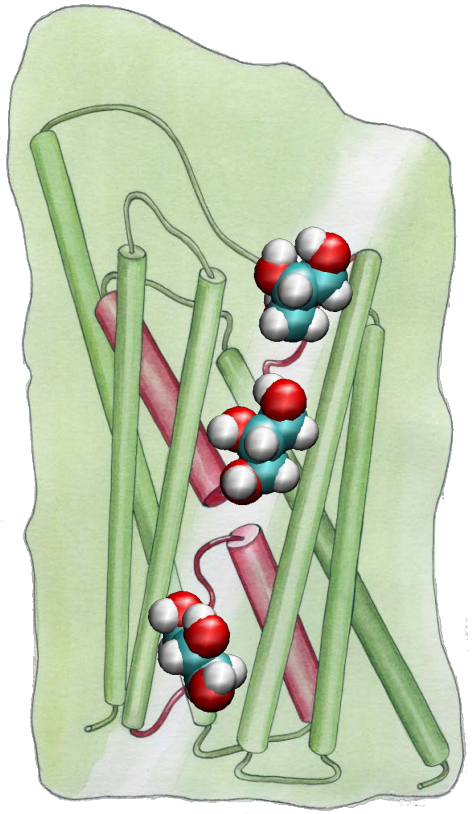
Constriction region



Selectivity filter

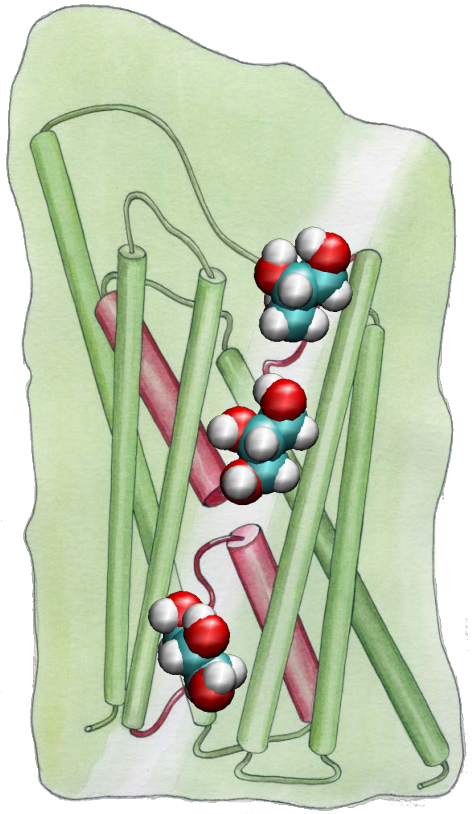
Channel Hydrogen Bonding Sites

```
…
{set frame 0}{frame < 100}{incr frame}{
    animate goto $frame
    set donor [atomselect top
    “name O N and within 2 of
    (resname GCL and name HO)”]
    lappend [$donor get index] list1
    set acceptor [atomselect top
    “resname GCL and name O and
    within 2 of (protein and name HN HO)”]
    lappend [$acceptor get index] list2
}
…
```



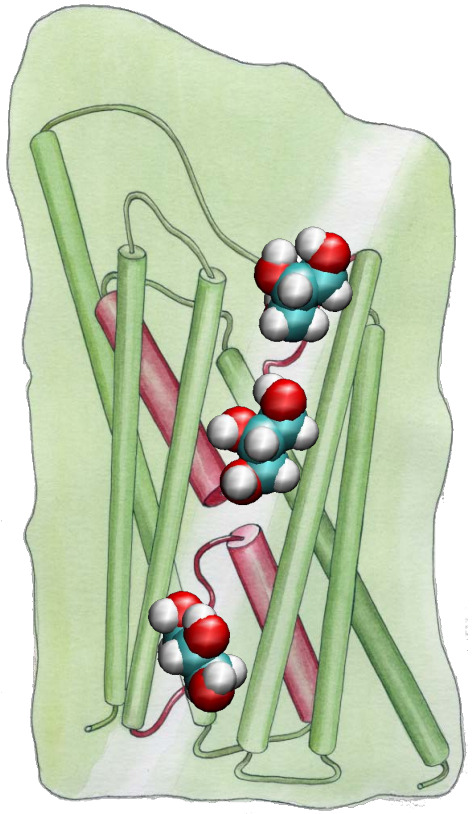
Channel Hydrogen Bonding Sites

GLN	41	OE1 NE2	LEU	197	O
TRP	48	O NE1	THR	198	O
GLY	64	O	GLY	199	O
ALA	65	O	PHE	200	O
HIS	66	O ND1	ALA	201	O
LEU	67	O	ASN	203	ND2
ASN	68	ND2			
ASP	130	OD1	LYS	33	HZ1 HZ3
GLY	133	O	GLN	41	HE21
SER	136	O	TRP	48	HE1
TYR	138	O	HIS	66	HD1
PRO	139	O N	ASN	68	HD22
ASN	140	OD1 ND2	TYR	138	HN
HIS	142	ND1	ASN	140	HN HD21 HD22
THR	167	OG1	HIS	142	HD1
GLY	195	O	GLY	199	HN
PRO	196	O	ASN	203	HN HD21HD22
			ARG	206	HE HH21HH22



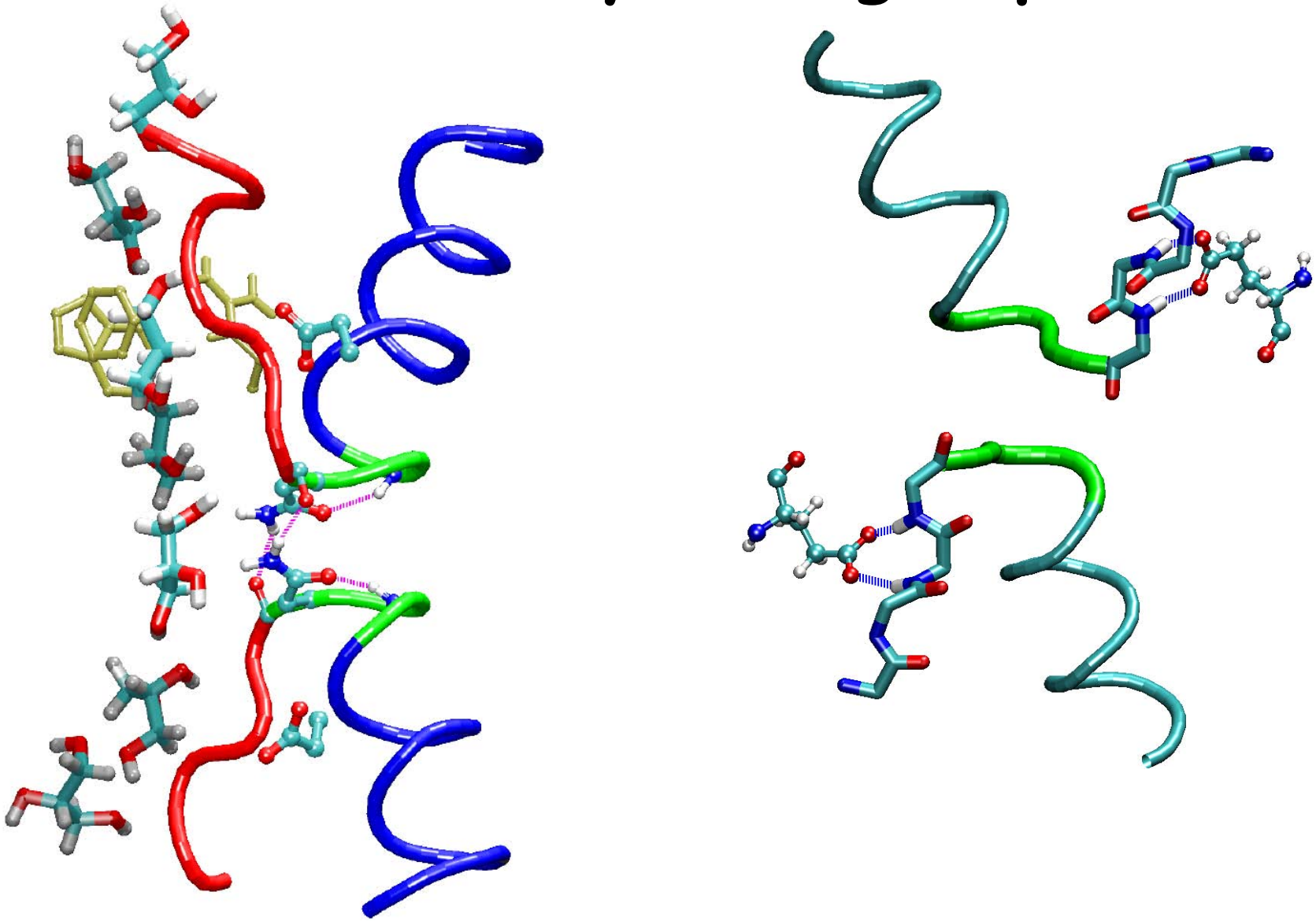
Channel Hydrogen Bonding Sites

GLN	41	OE1 NE2	LEU	197	O
TRP	48	O NE1	THR	198	O
GLY	64	O	GLY	199	O
ALA	65	O	PHE	200	O
HIS	66	O ND1	ALA	201	O
LEU	67	O	ASN	203	ND2
ASN	68	ND2			
ASP	130	OD1	LYS	33	HZ1 HZ3
GLY	133	O	GLN	41	HE21
SER	136	O	TRP	48	HE1
TYR	138	O	HIS	66	HD1
PRO	139	O N	ASN	68	HD22
ASN	140	OD1 ND2	TYR	138	HN
HIS	142	ND1	ASN	140	HN HD21 HD22
THR	167	OG1	HIS	142	HD1
GLY	195	O	GLY	199	HN
PRO	196	O	ASN	203	HN HD21HD22
			ARG	206	HE HH21HH22



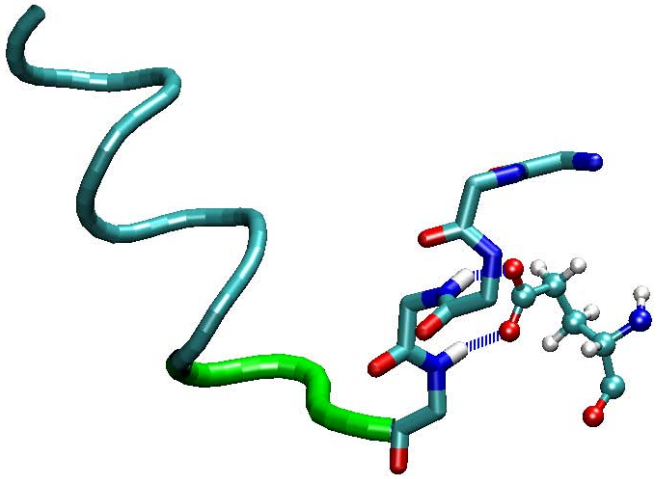
The Substrate Pathway

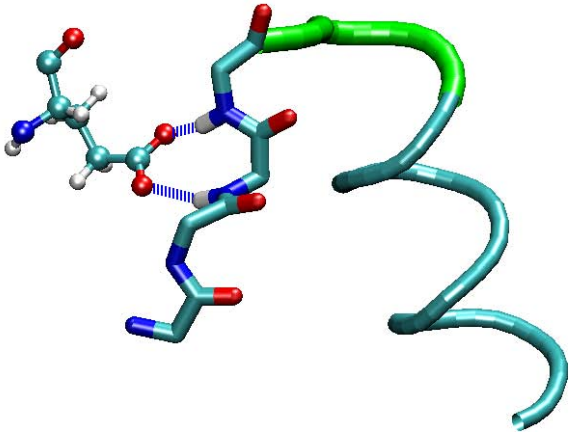
is formed by C=O groups



The Substrate Pathway is formed by C=O groups

Non-helical motifs are stabilized by two glutamate residues.



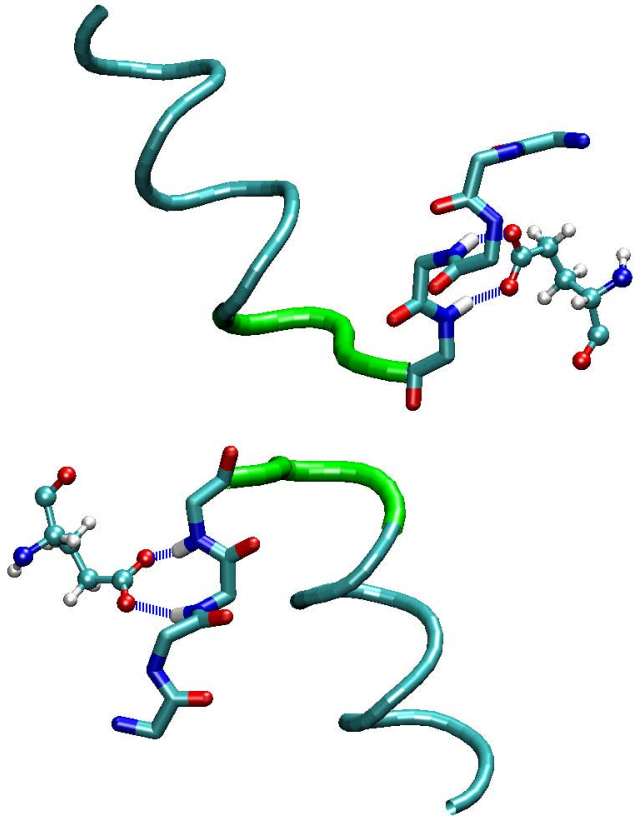


N 

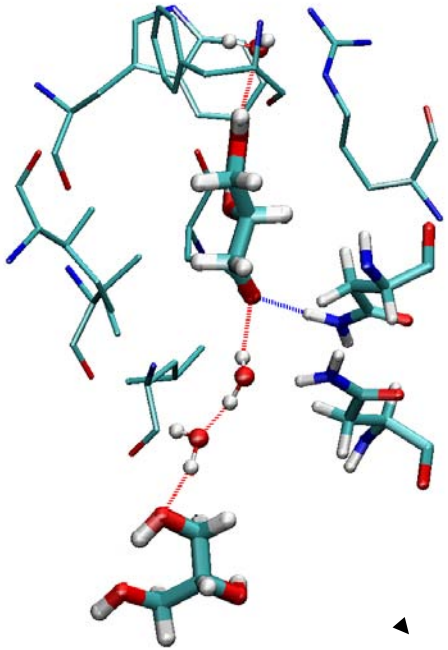
C

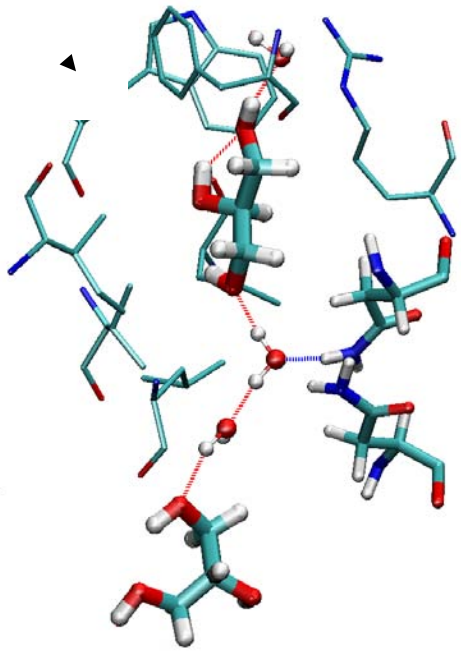
Conservation of Glutamate Residue in Human Aquaporins

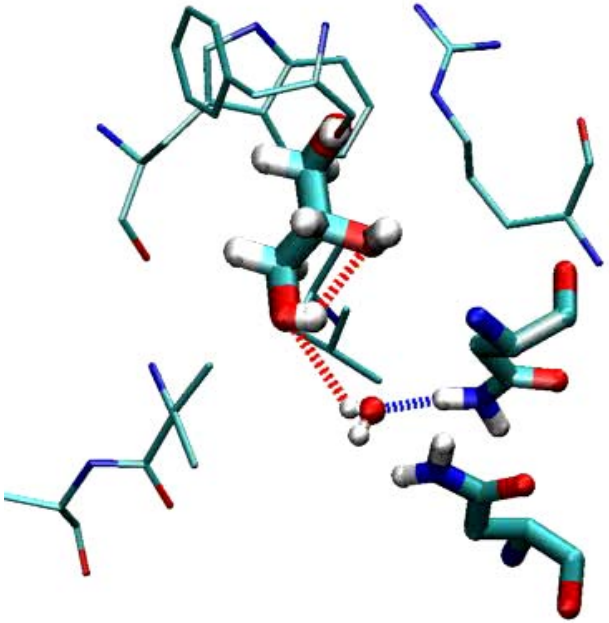
```
                                .   :   *   :   :
AQP0_HUMAN  ---LNTLHPAVSVGQATTVEIFLTLQFVLCIFATYDE-RRNGQLG
AQP1_HUMAN  ---RNDLADGVNSGQGLGIEIIGTLQLVLCVLATTDR-RRRDLGG
AQP2_HUMAN  ---VNALSNSTTAGQAVTVELFLTLQLVLCIFASTDE-RRGENPG
AQP3_HUMAN  GIFATYPSGHLDMINGFFDQFIGTASLIVCVLAIVDPYNNPVPRG
AQP4_HUMAN  ---VTMVHGNLTAGHGLLVELIITFQLVFTIFASCDS-KRTDVTG
AQP5_HUMAN  ---VNALNNNTTQGQAMVVELILTFQLALCIFASTDS-RRTSPVG
AQP6_HUMAN  ---INVVRNSVSTGQAVAVELLLTLQLVLCVFASTDS-RQTS--G
AQP7_HUMAN  GIFATYLPDHMTLWRGFLNEAWLTGMLQLCLFAITDQENNPALPG
AQP8_HUMAN  -AAFVTVQEQGQVAGALVAEIILTTLLALAVCMGAIN--EKTKGP
AQP9_HUMAN  HIFATYPAPYLSLANAFADQVVATMILLIIVFAIFDSRNLGAPRG
GLPF_ECOLI  GTFSTYPNPHINFVQAFAVEMVITAILMGLILALTDDGNGVPRGP
      ruler ...180.......190.......200.......210.......22
```



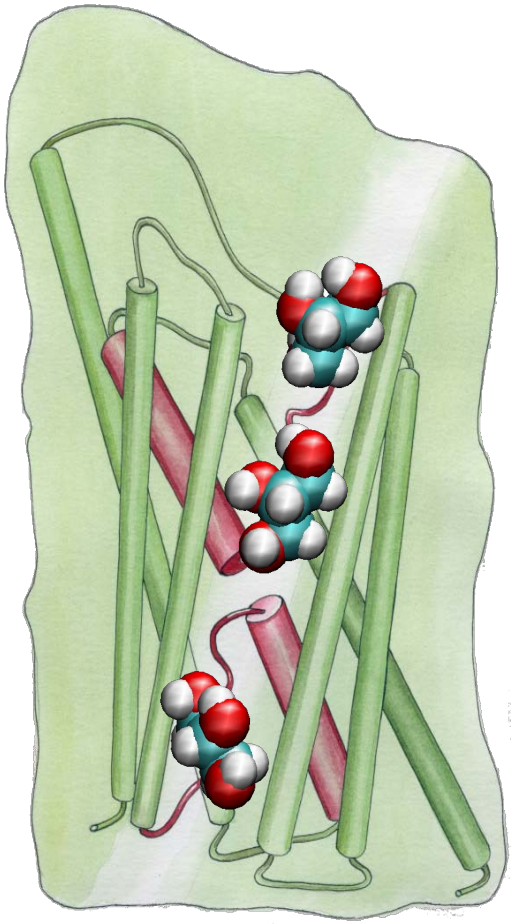
Glycerol – water competition for hydrogen bonding sites

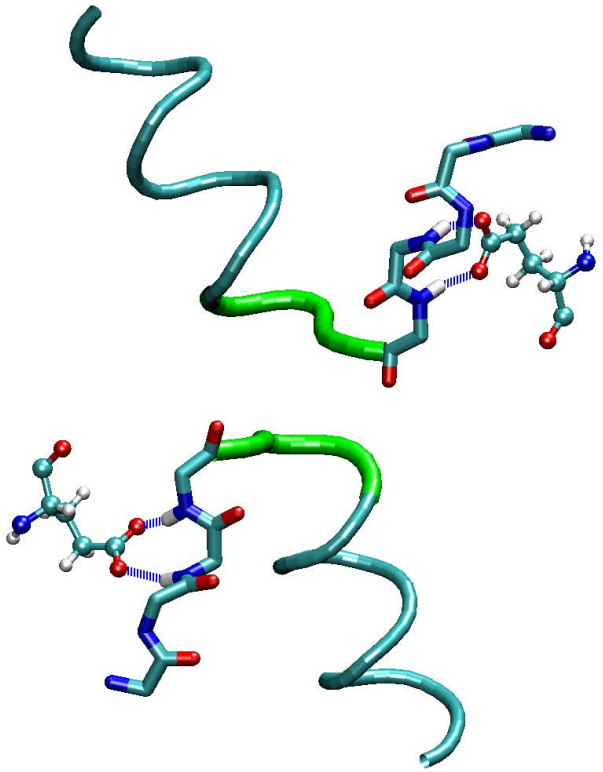




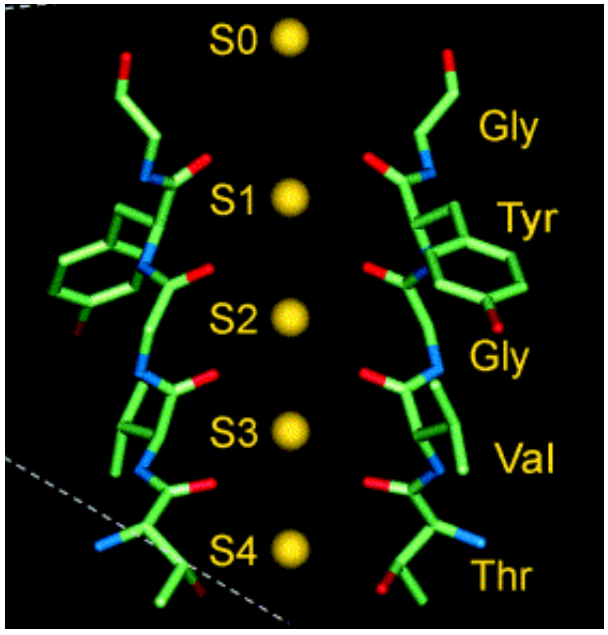


Revealing the Functional Role of Reentrant Loops



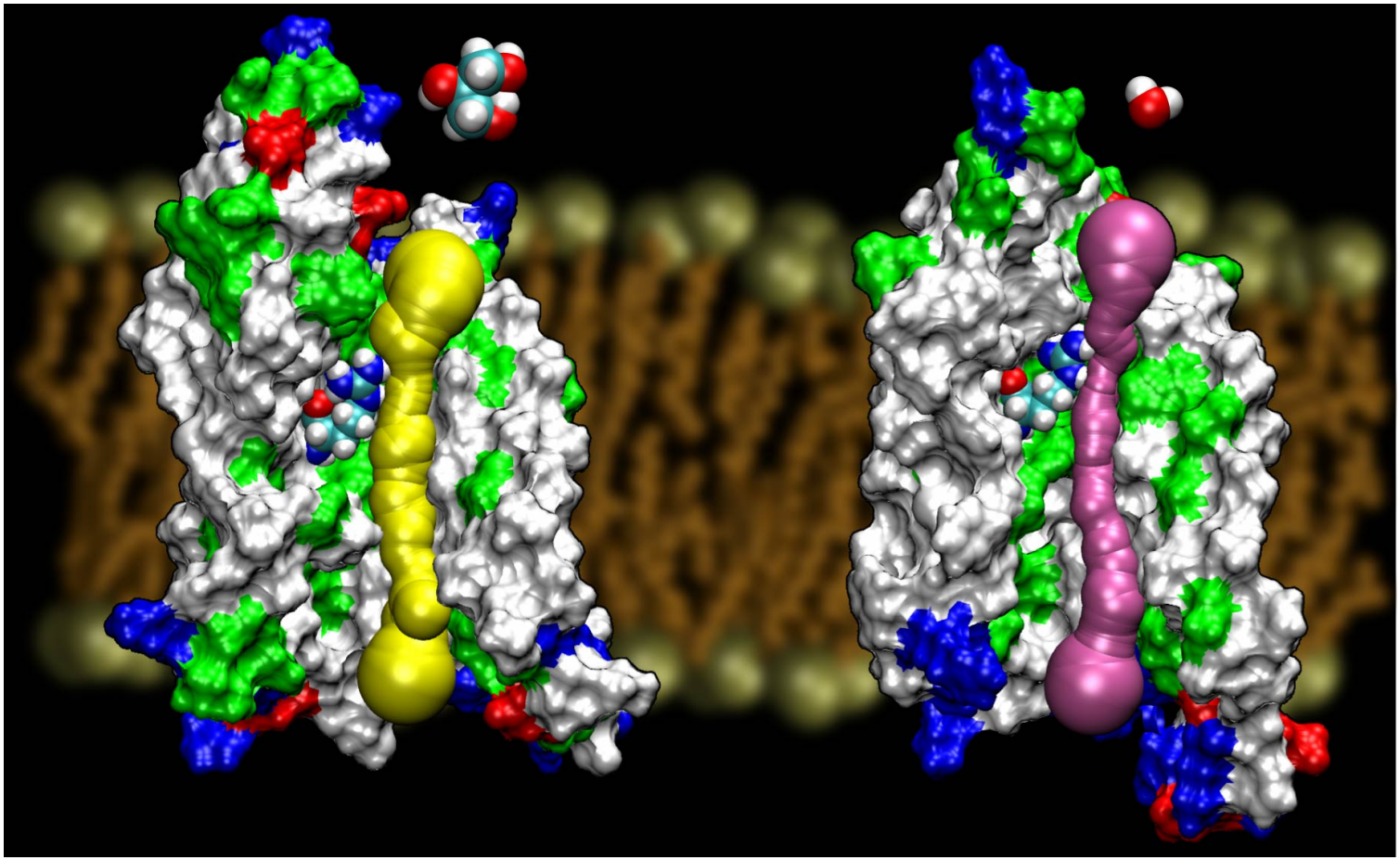


Potassium channel



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 is a non-equilibrium method by nature

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

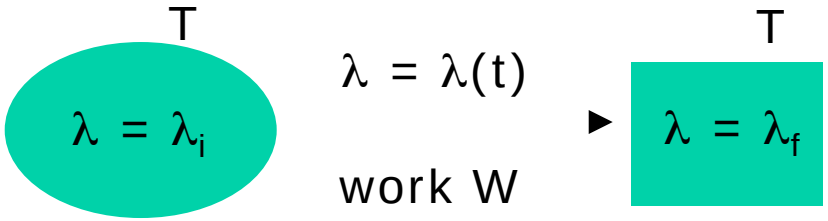
- The system will be driven, however, away from equilibrium, resulting in problems in describing the energy landscape associated with the event of interest.

Second law of thermodynamics

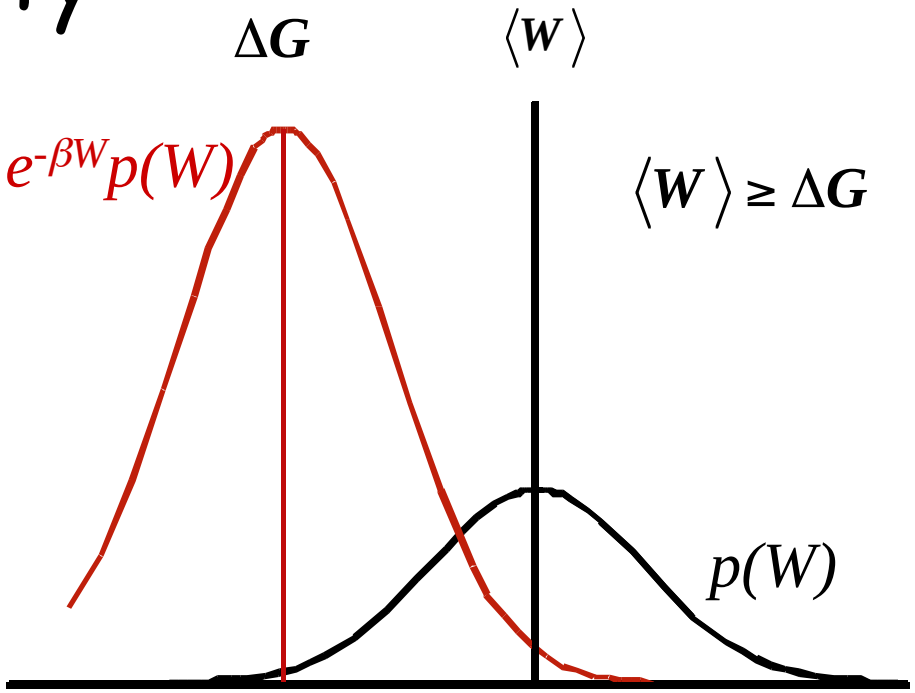
$$W \geq \Delta G$$

Jarzynski's Equality

Transition between two equilibrium states



$$\Delta G = G_f - G_i$$



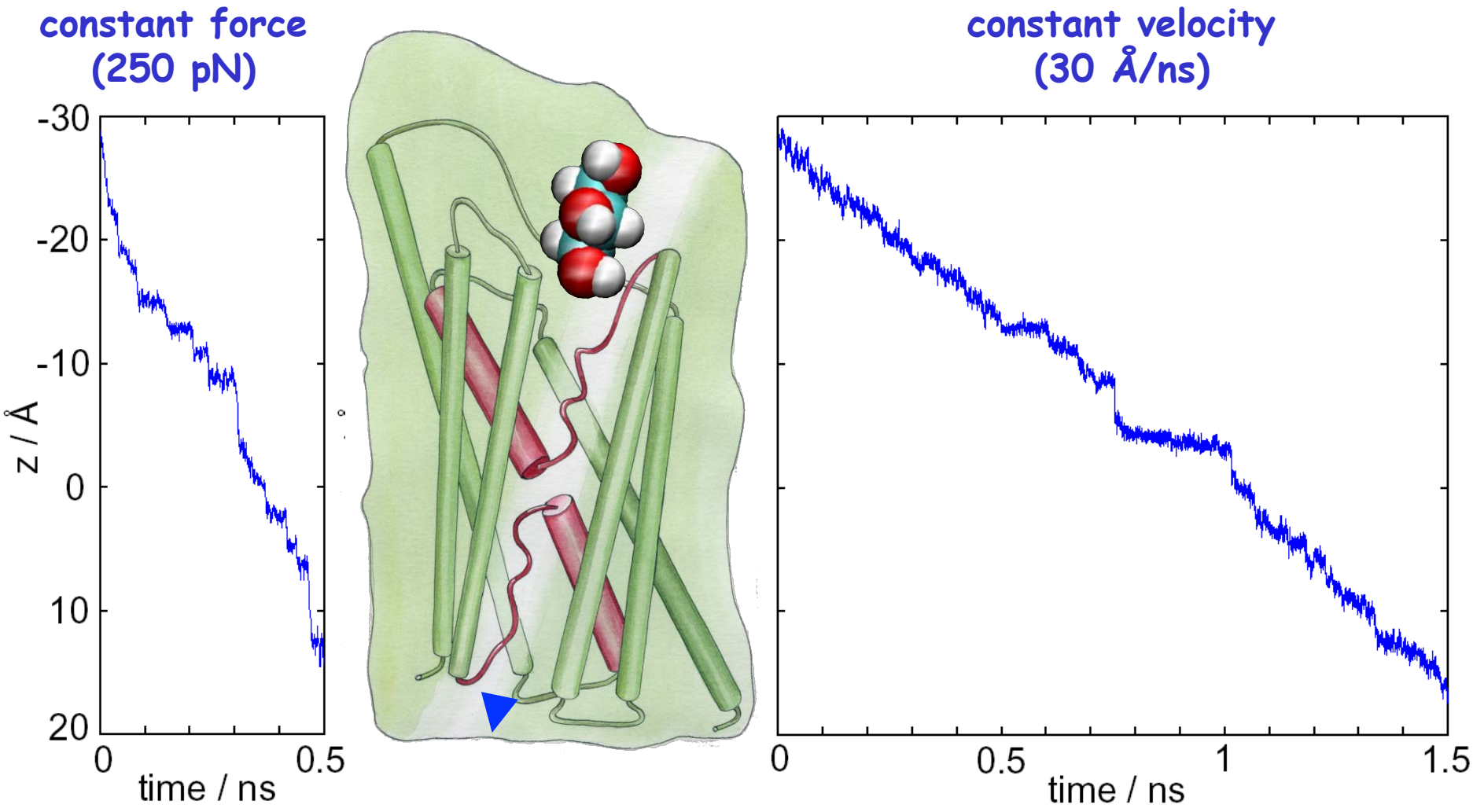
C. Jarzynski, *Phys. Rev. Lett.*, 78, 2690 (1997)
C. Jarzynski, *Phys. Rev. E*, 56, 5018 (1997)

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

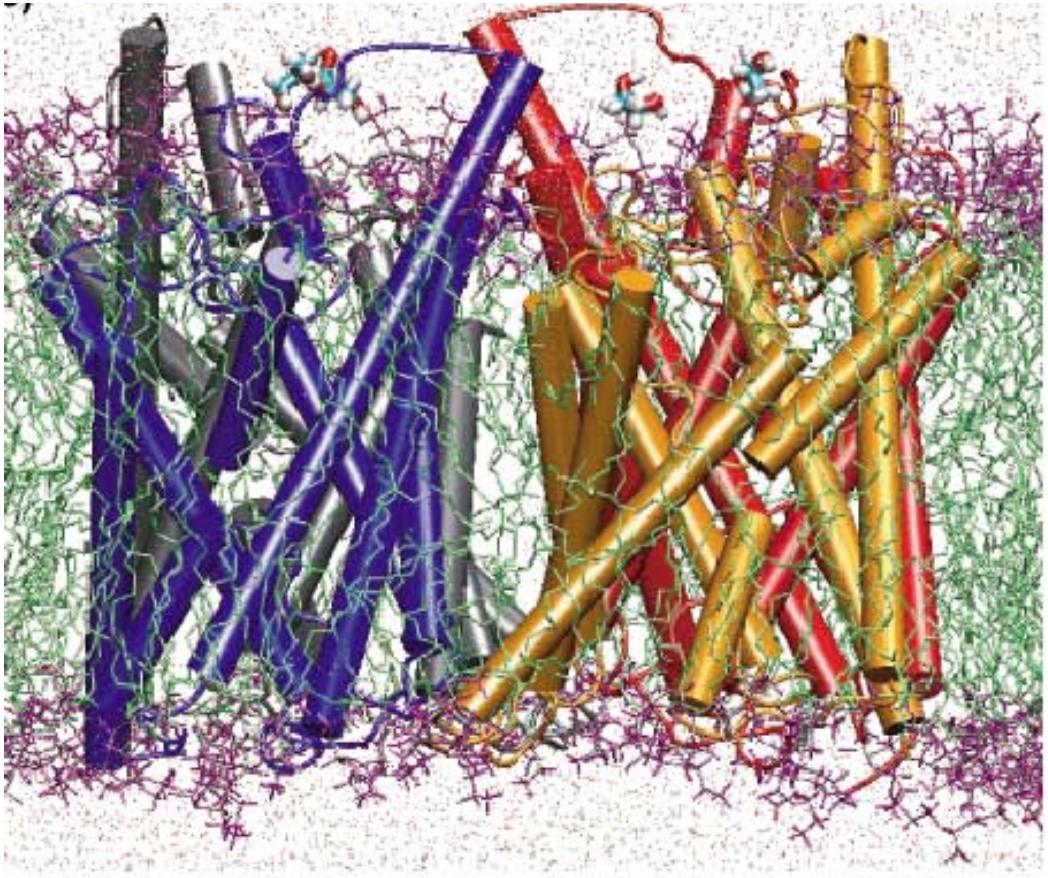
$$\beta = 1/k_B T$$

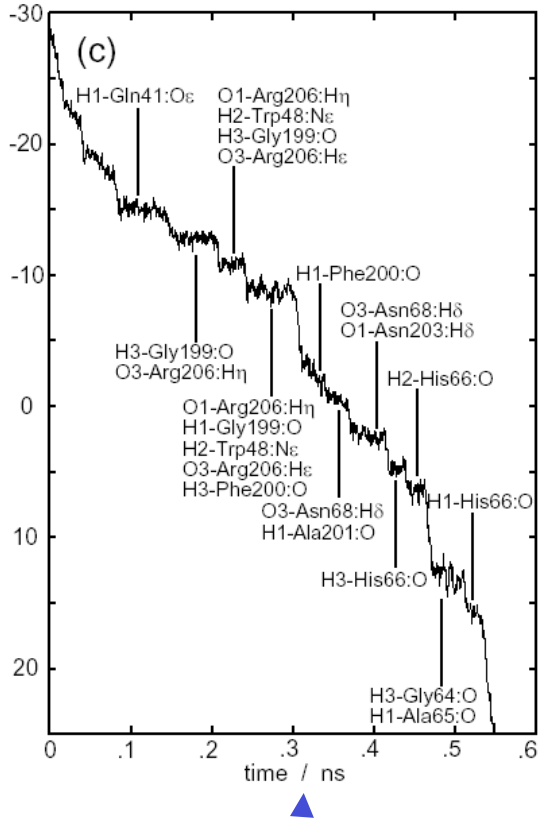
In principle, it is possible to obtain free energy surfaces from *repeated non-equilibrium* experiments.

Steered Molecular Dynamics



SMD Simulation of Glycerol Passage





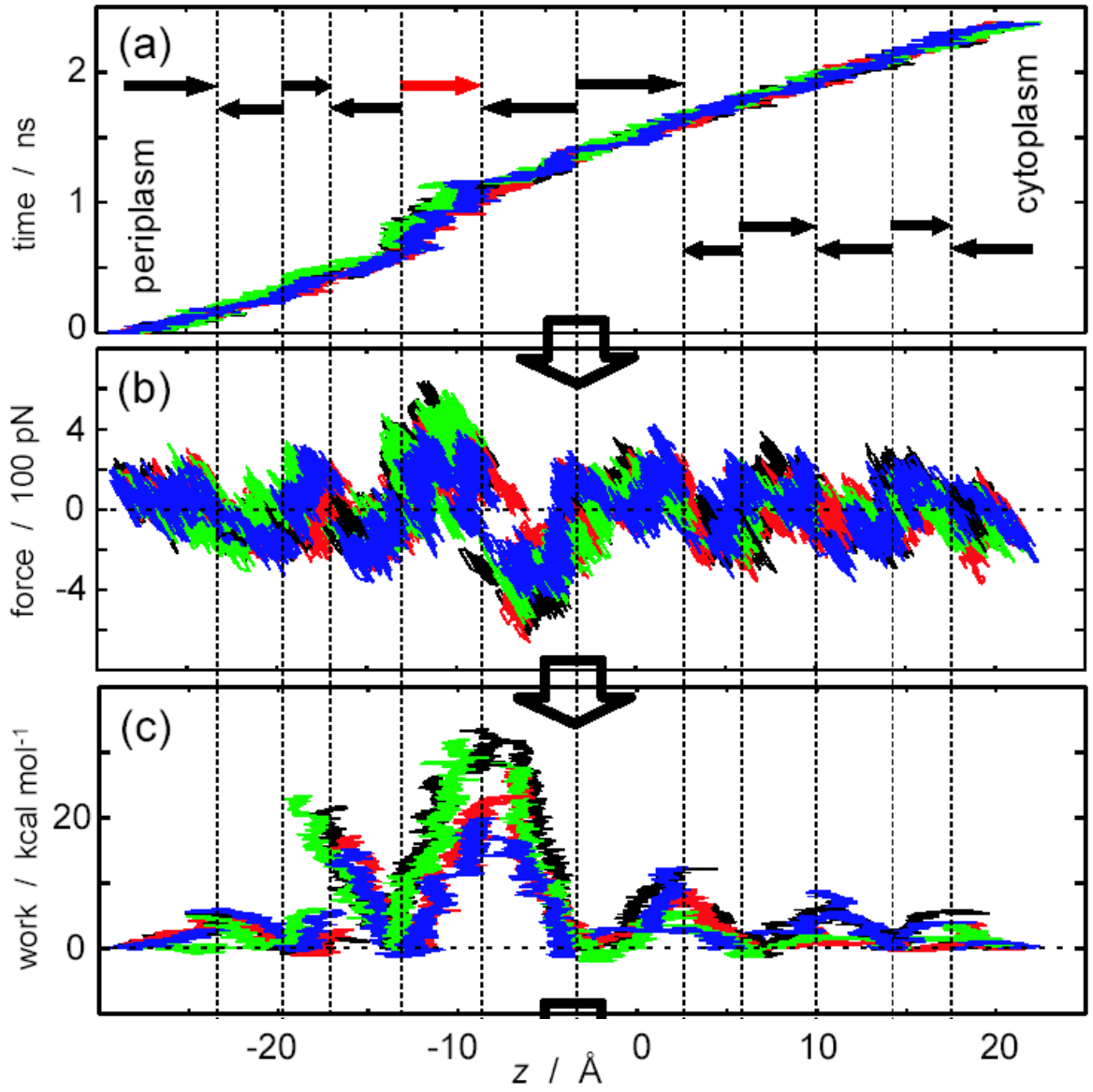
Trajectory of glycerol pulled by **constant force**

Constructing the Potential of Mean Force

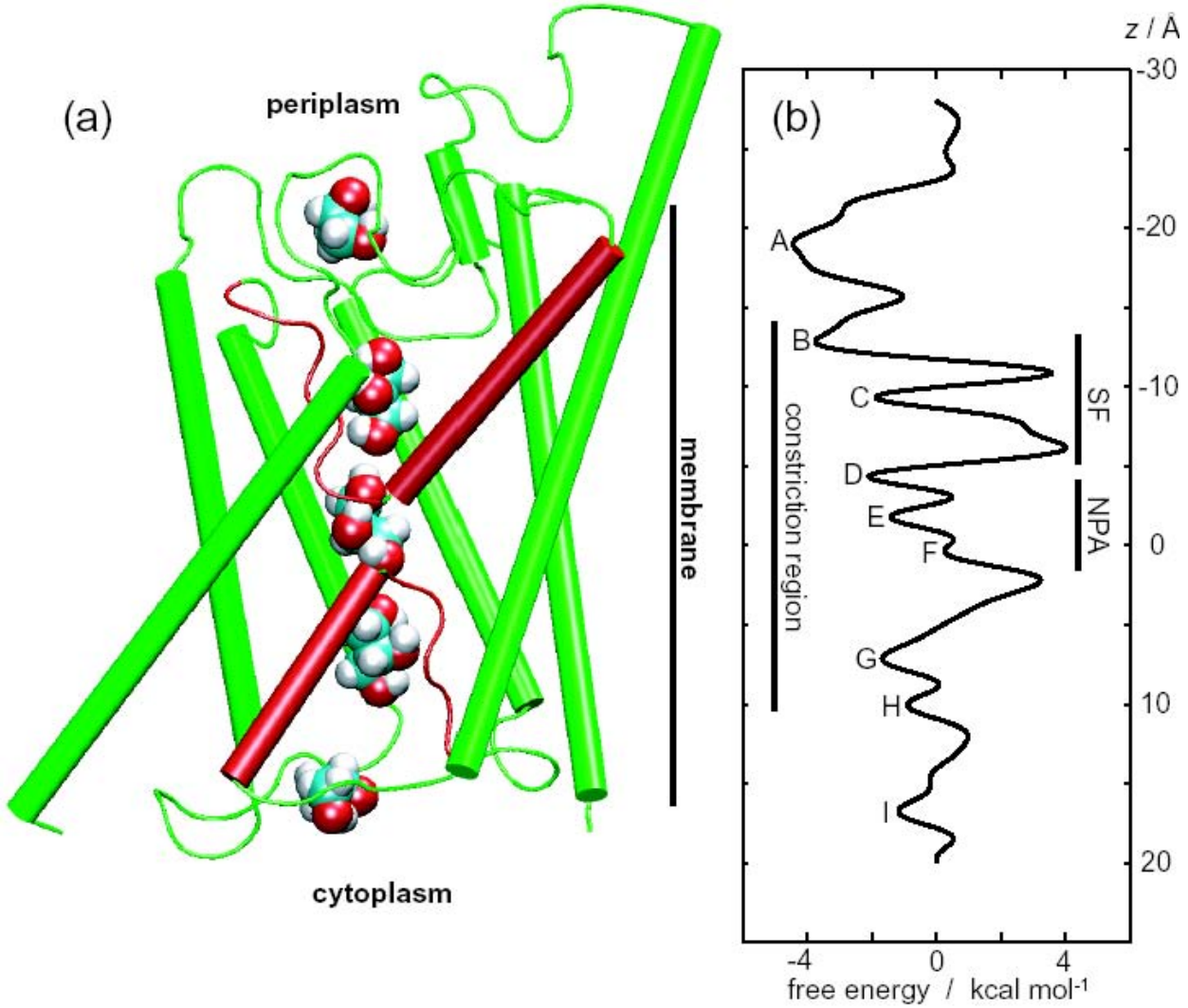
4 trajectories
v **= 0.03, 0.015 Å/ps**
k **= 150 pN/Å**

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

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



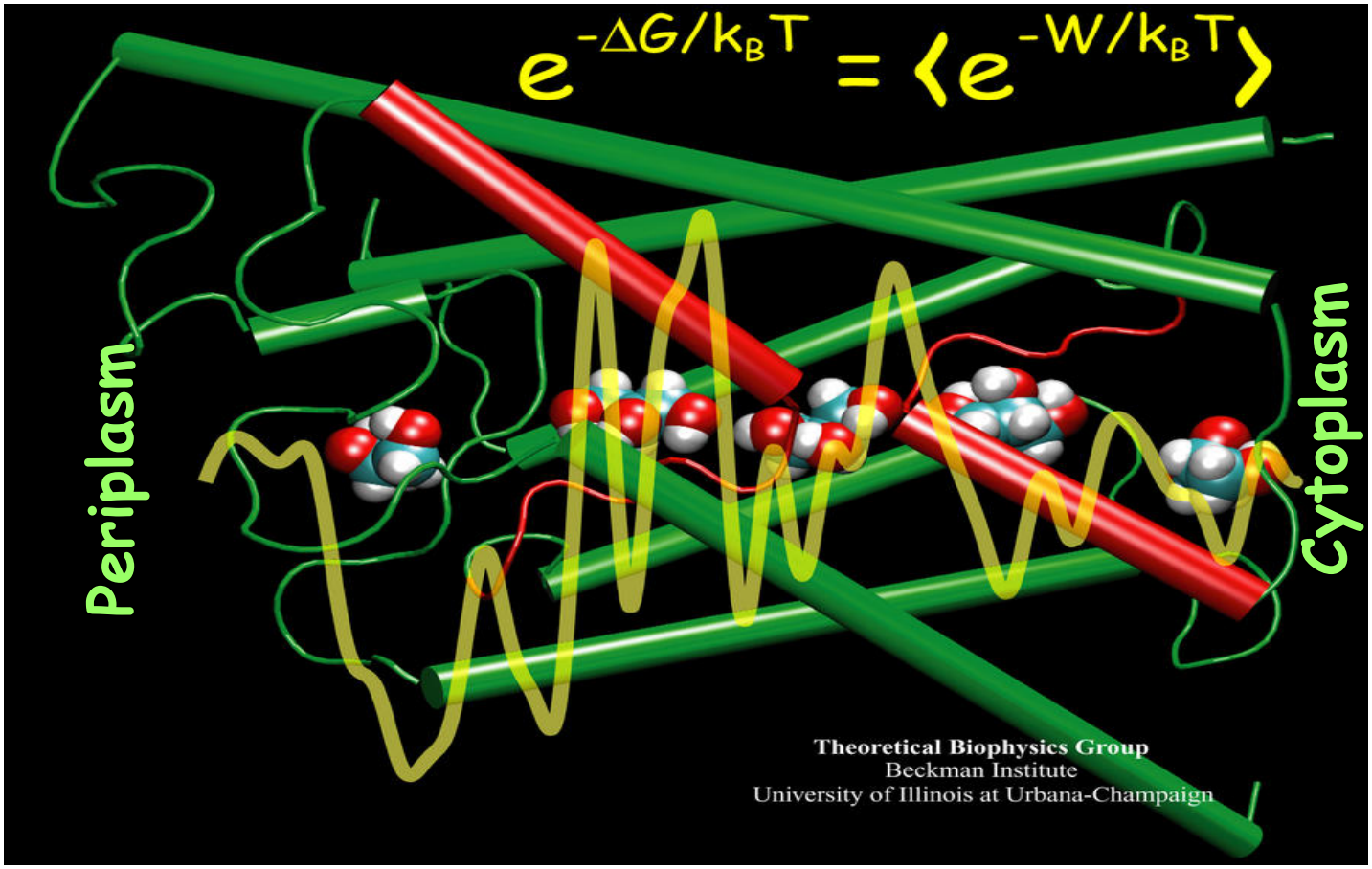
Features of the Potential of Mean Force



- 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.

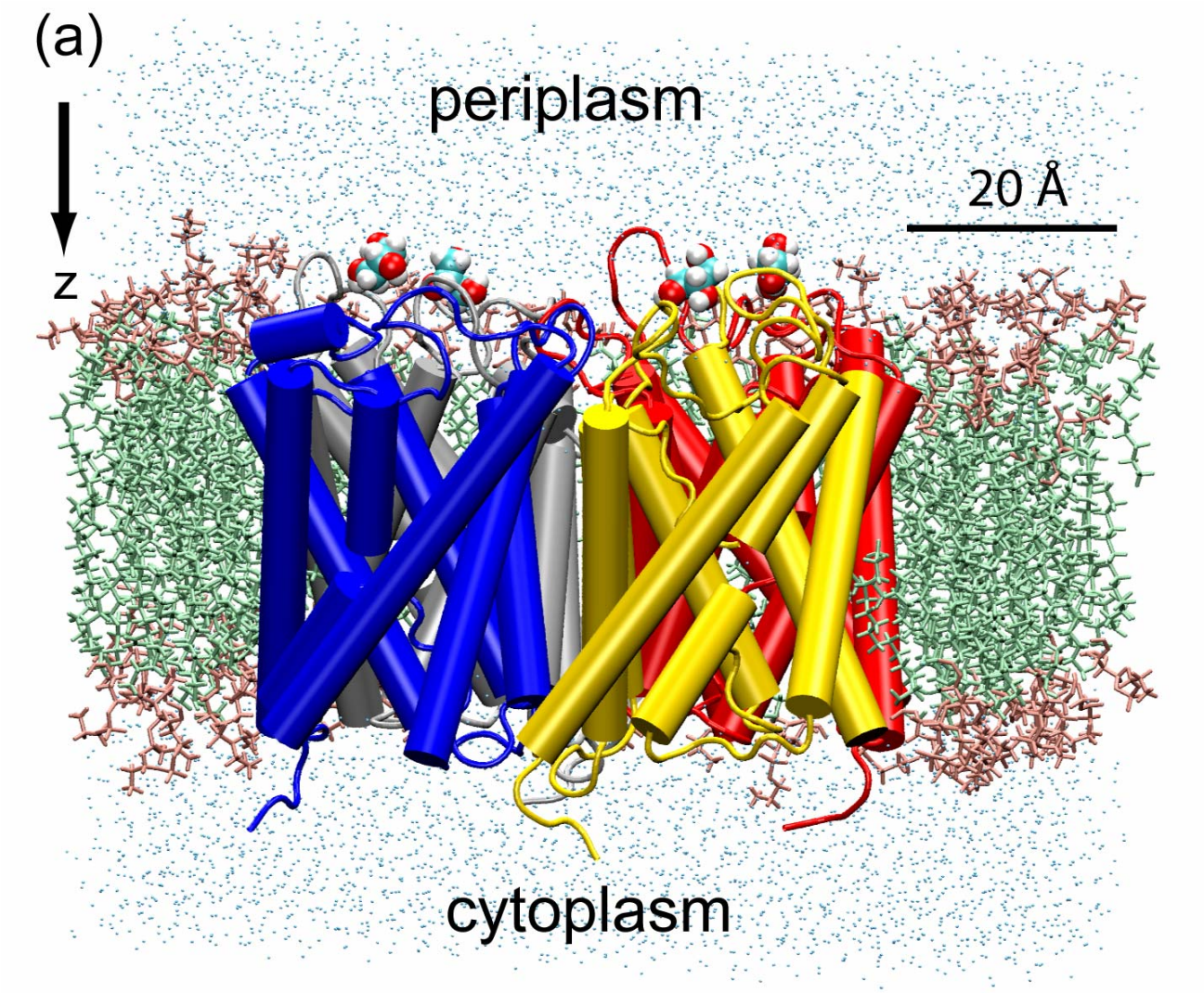
Features of the Potential of Mean Force



Asymmetric Profile in the Vestibules

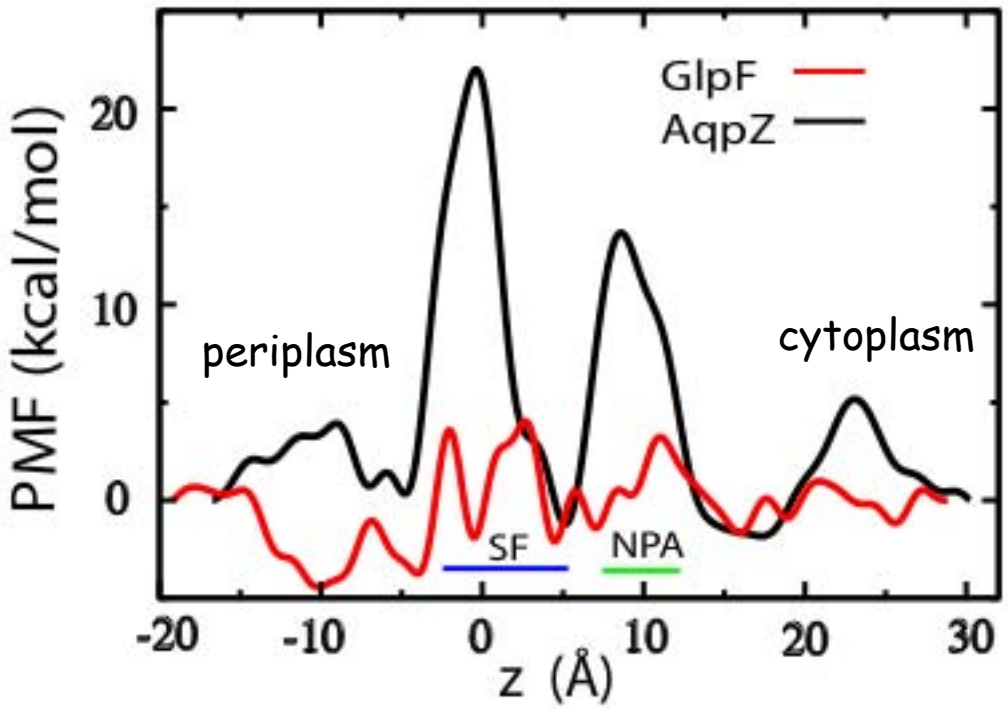
Jensen et al., *PNAS*, 99:6731-6736, 2002.

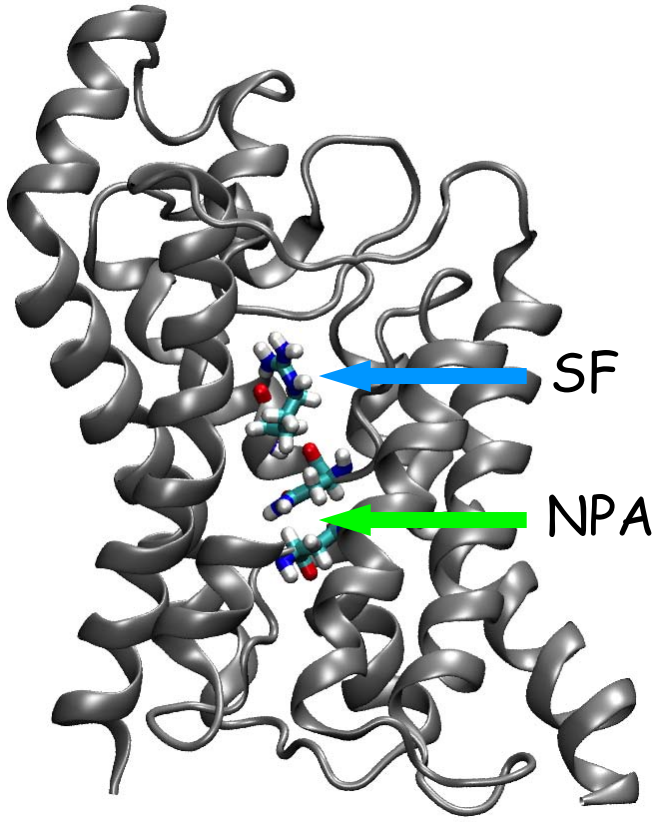
Artificial induction of glycerol conduction through AqpZ



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

Three fold higher barriers

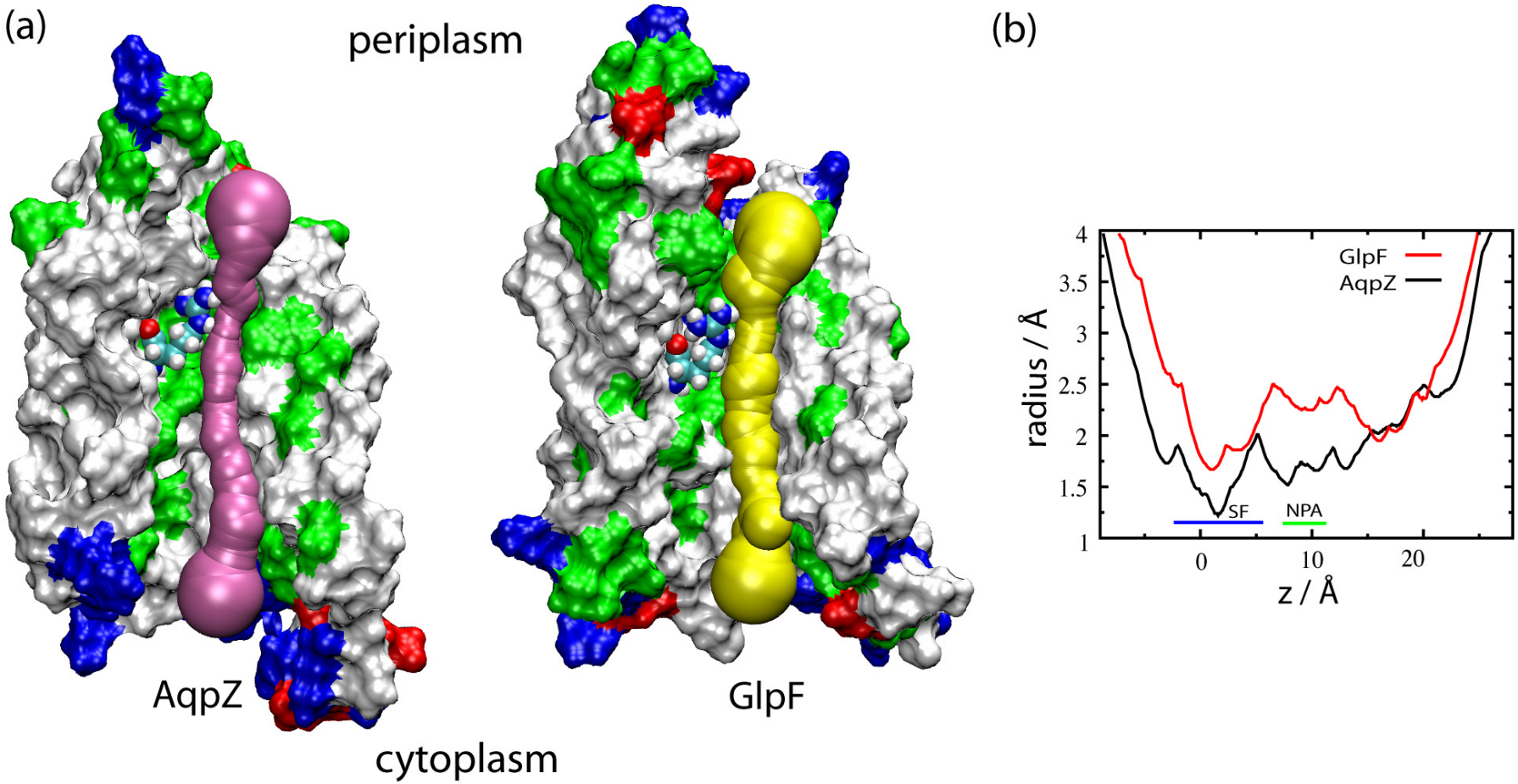




AqpZ 22.8 kcal/mol
GlpF 7.3 kcal/mol

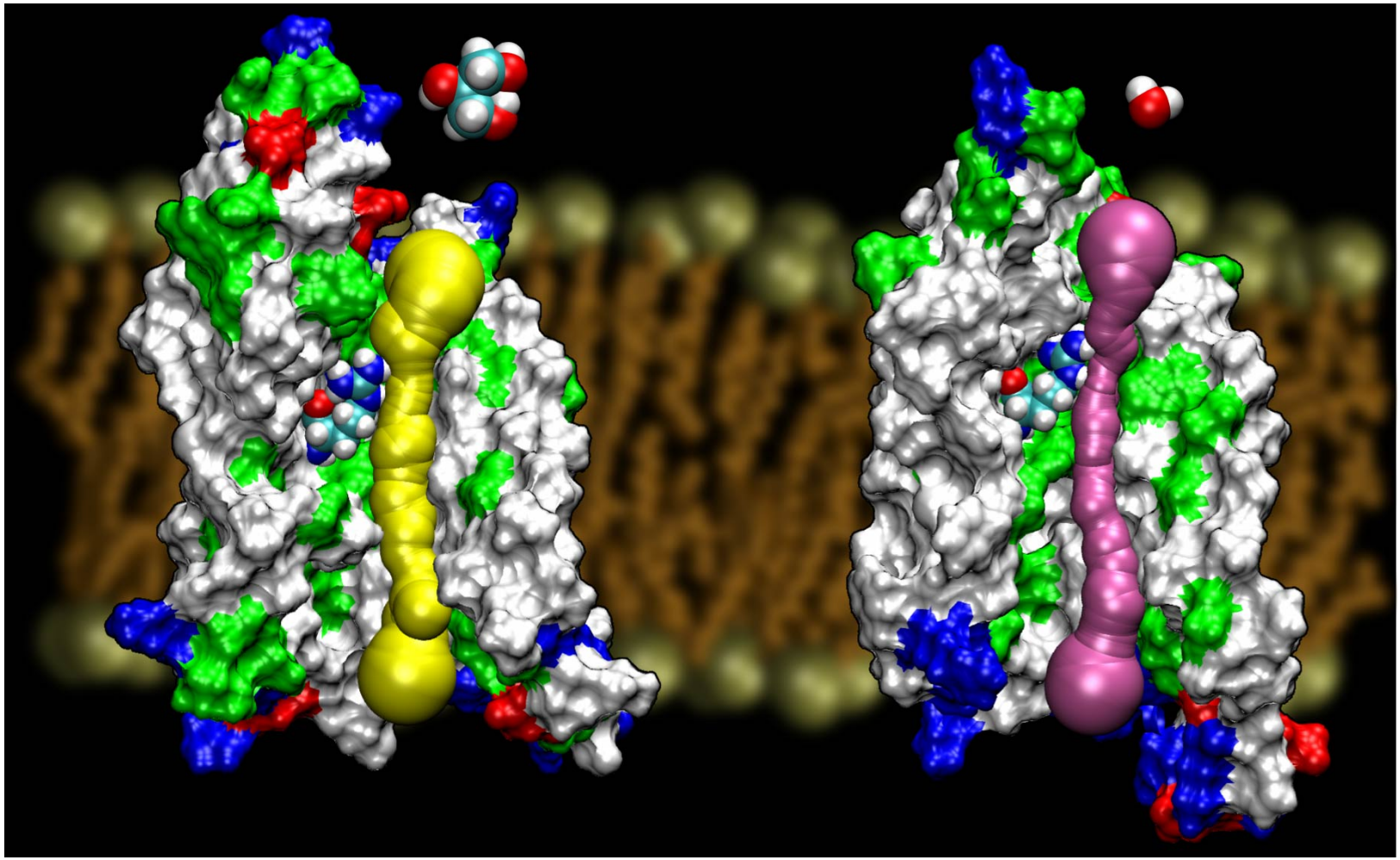
Y. Wang, K. Schulten, and E. Tajkhorshid ***Structure*** 13, 1107 ***(***2005)

Could it be simply the size?



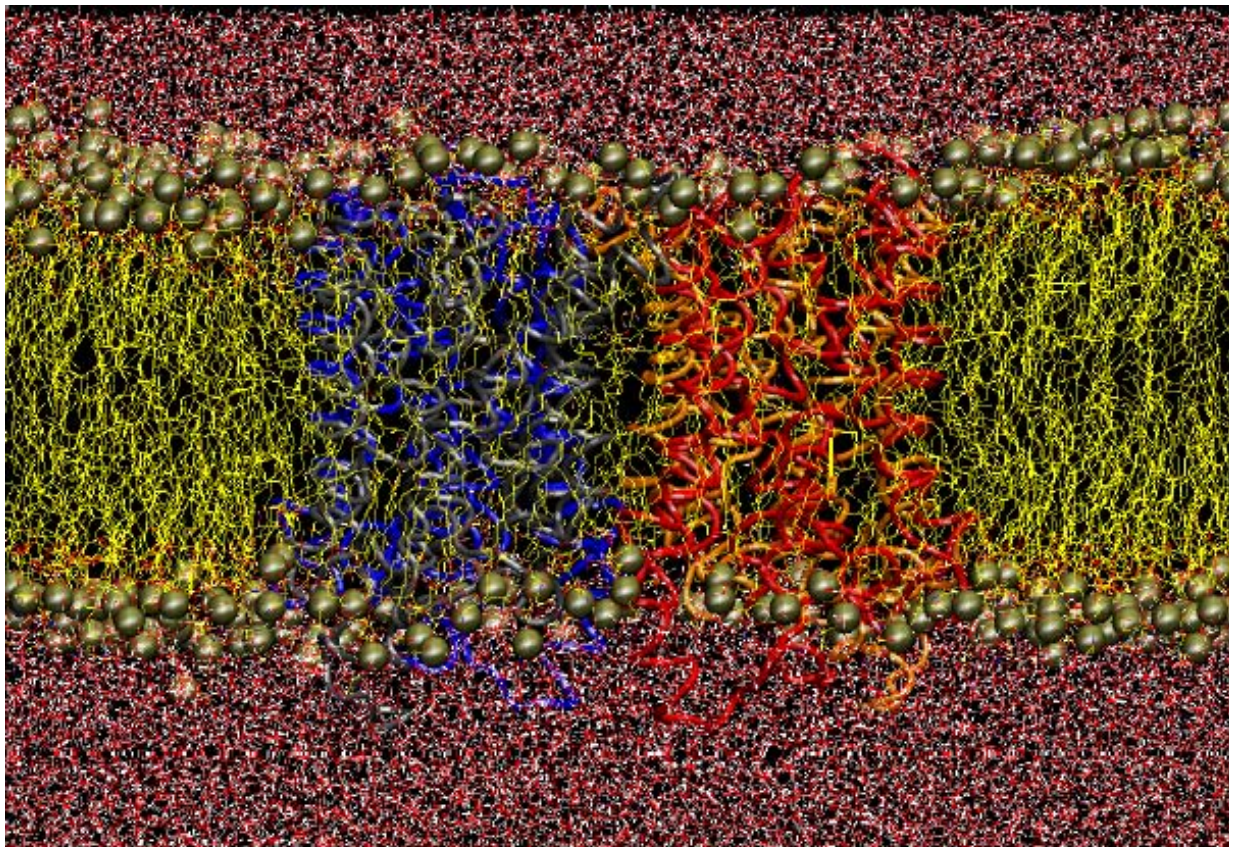
Y. Wang, K. Schulten, and E. Tajkhorshid ***Structure*** 13, 1107 ***(***2005)

It is probably just the size that matters!



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

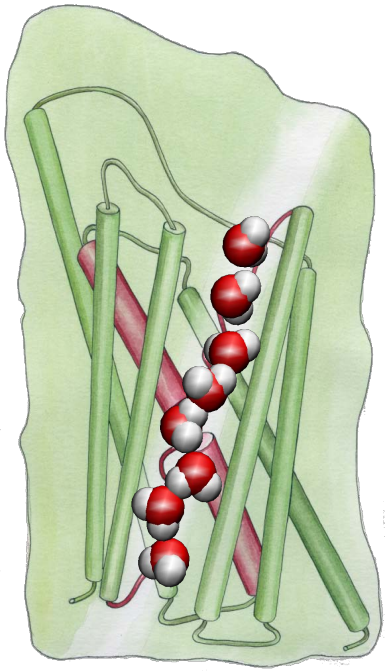
Water permeation



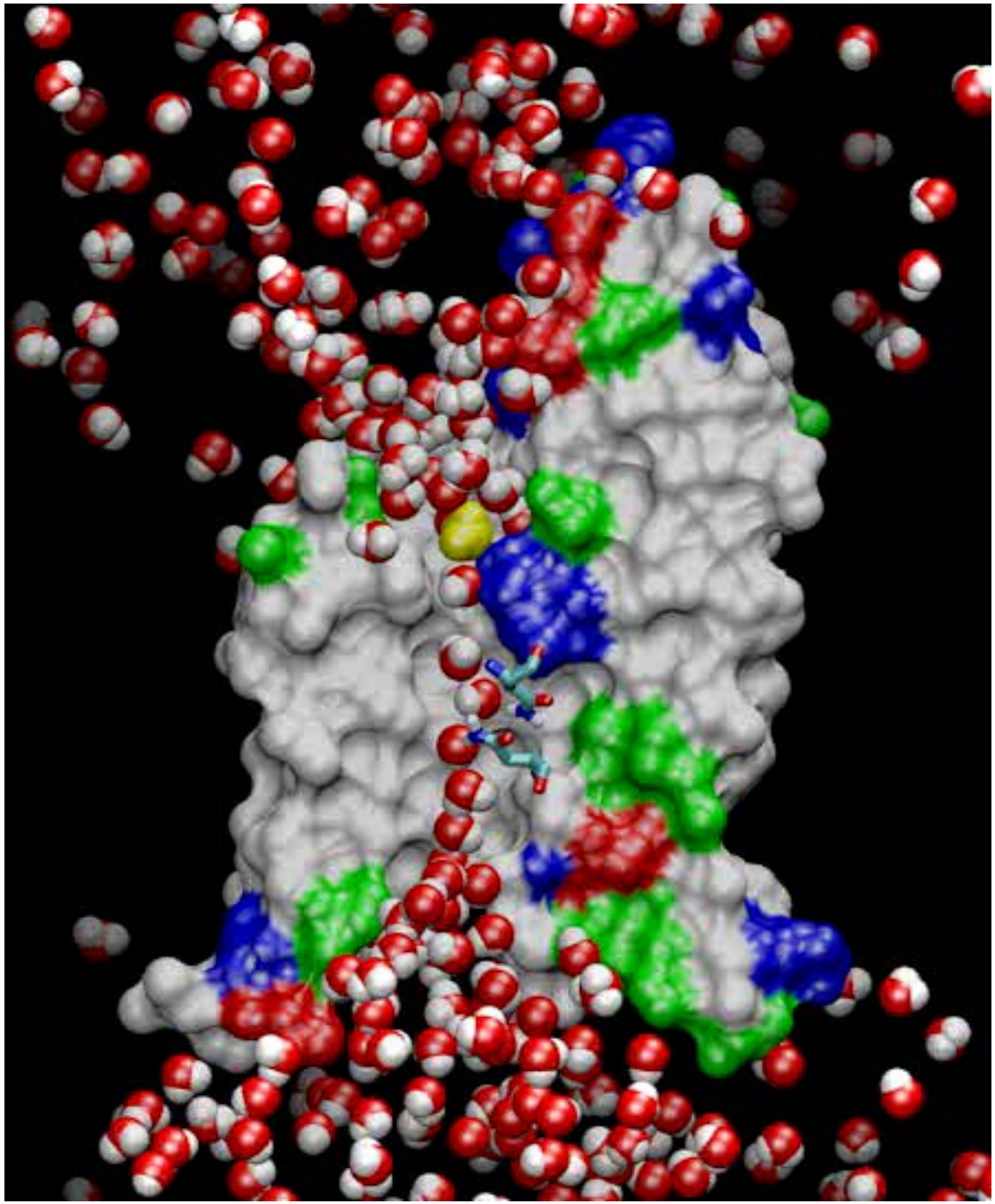
18 water conducted
In 4 monomers in 4 ns
1.125 water/monomer/ns

Exp. = ~ 1-2 /ns

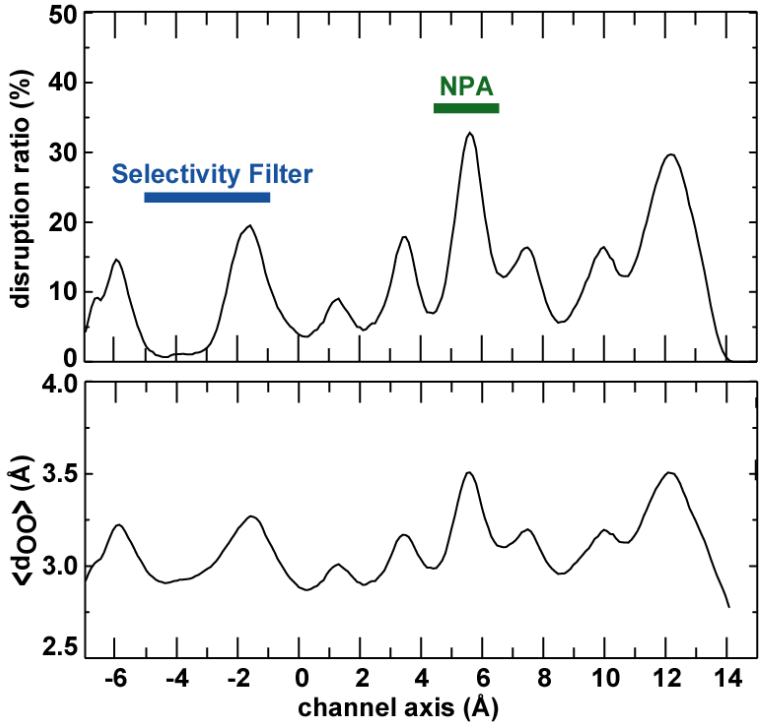
5 nanosecond Simulation

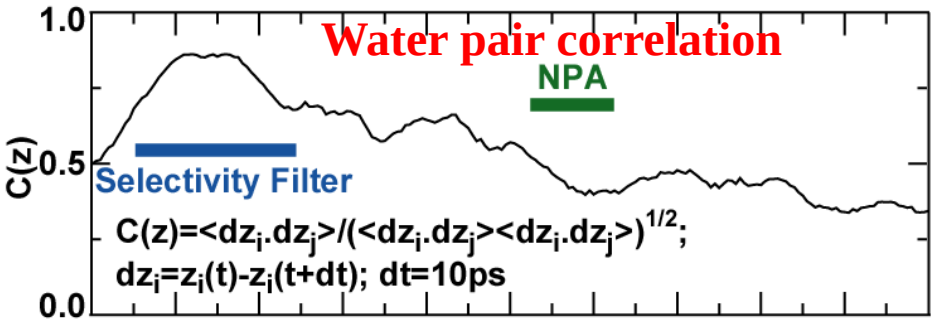


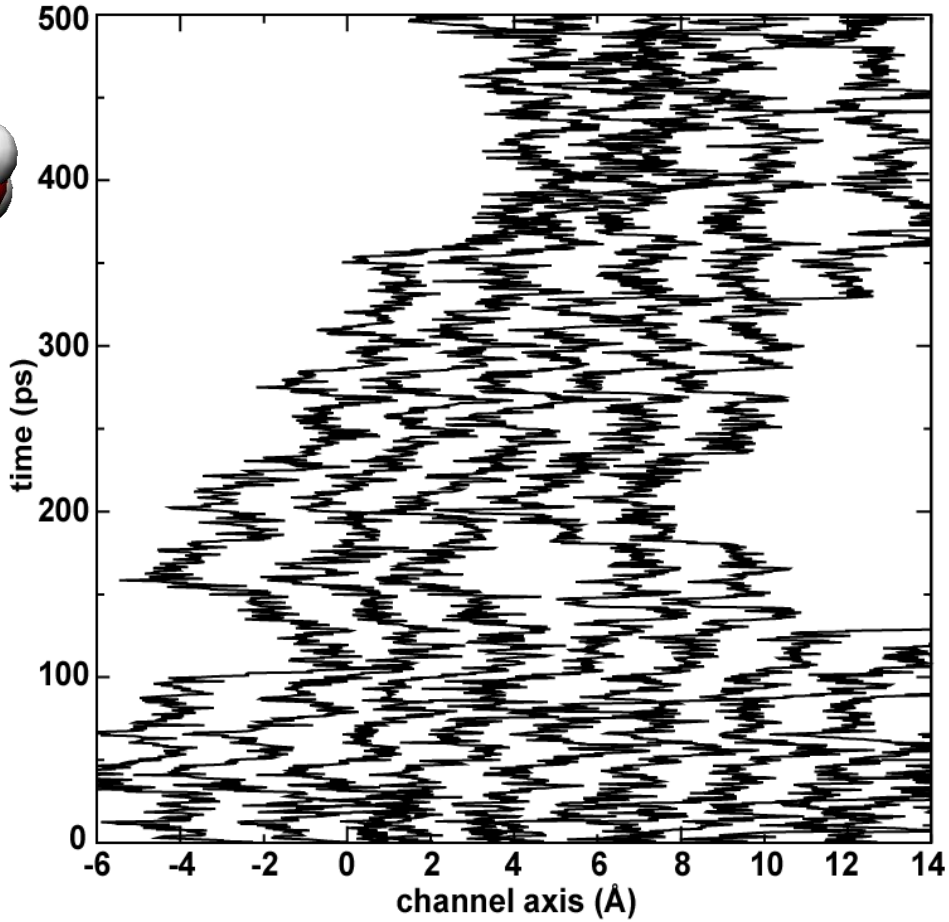
7-8 water molecules in each channel

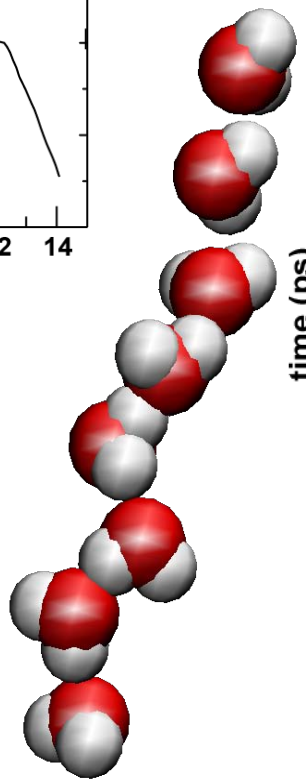


Correlated Motion of Water in the Channel



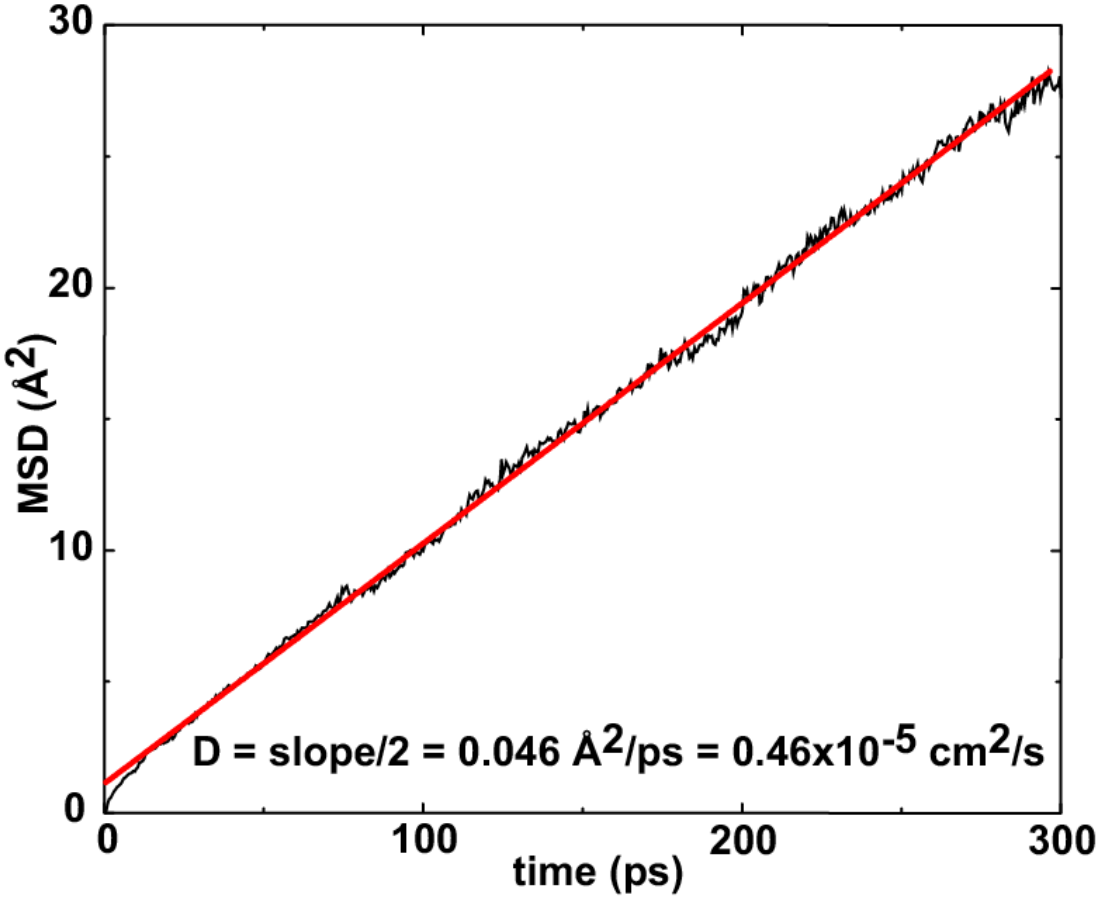


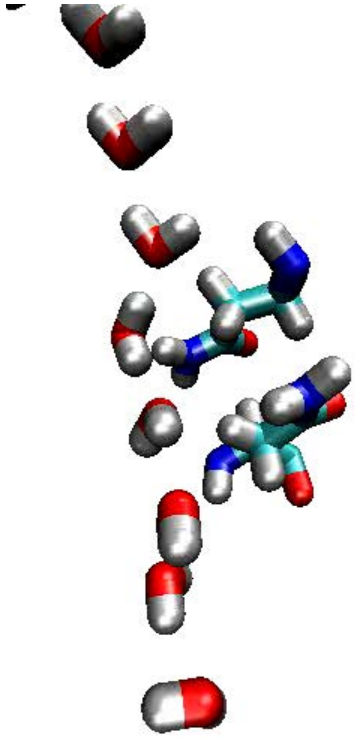




The single file of water molecules is maintained.

Diffusion of Water in the channel

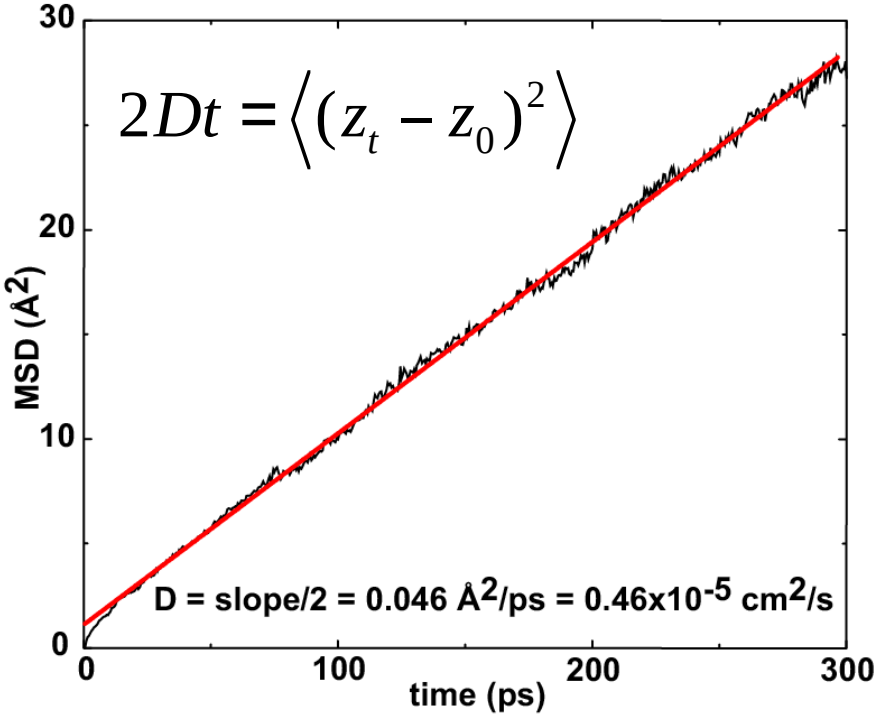


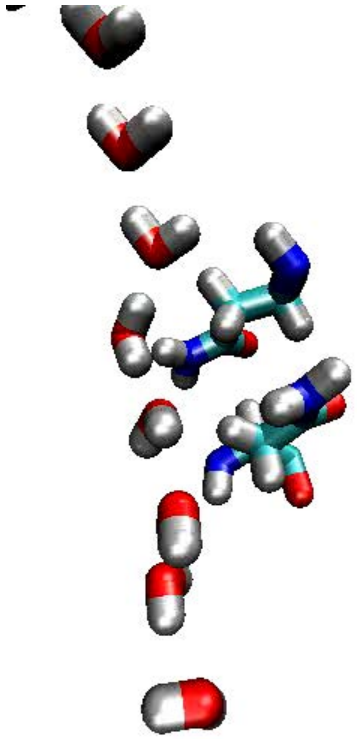


One dimensional diffusion: $2Dt = \left\langle (z_t - z_0)^2 \right\rangle$

Experimental value for AQP1: 0.4-0.8 e-5

Diffusion of Water in the channel



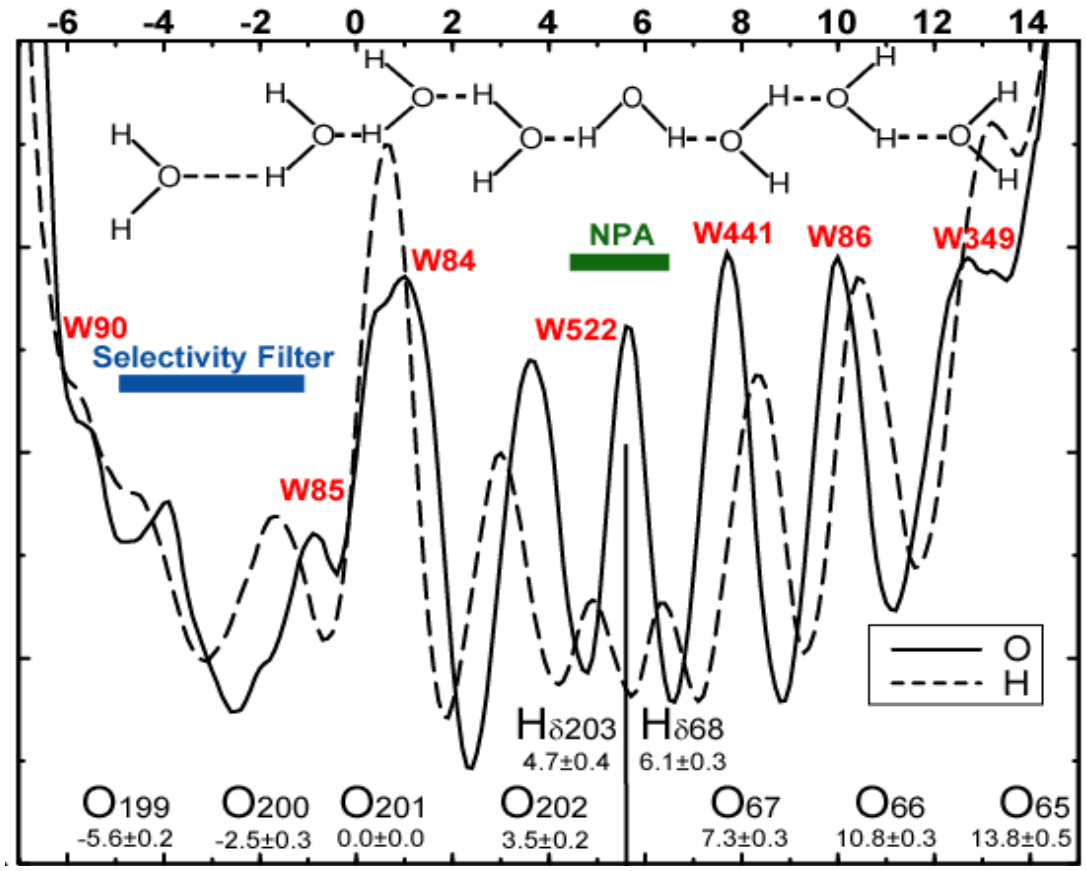


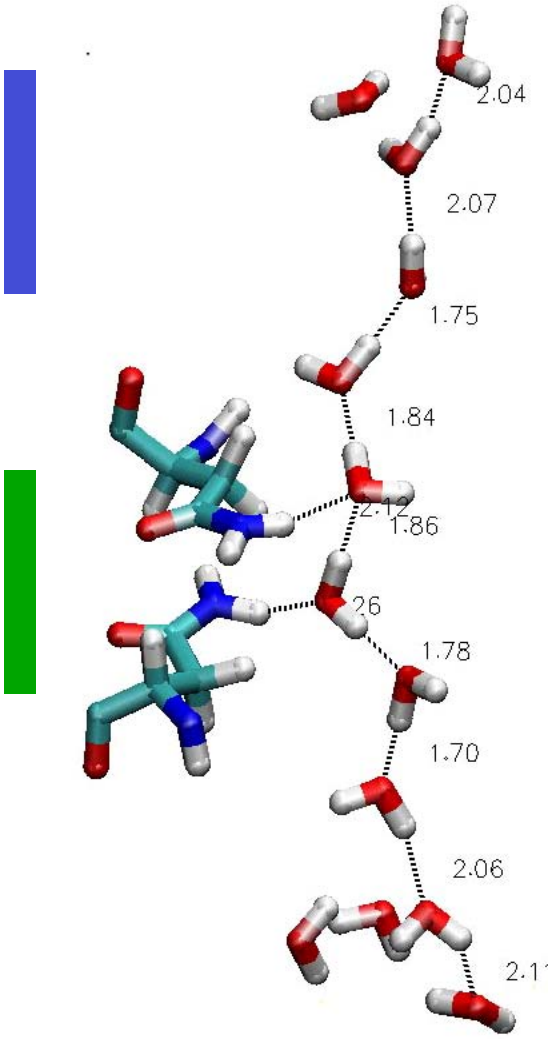
0 1 2 3 4

Time (ns)

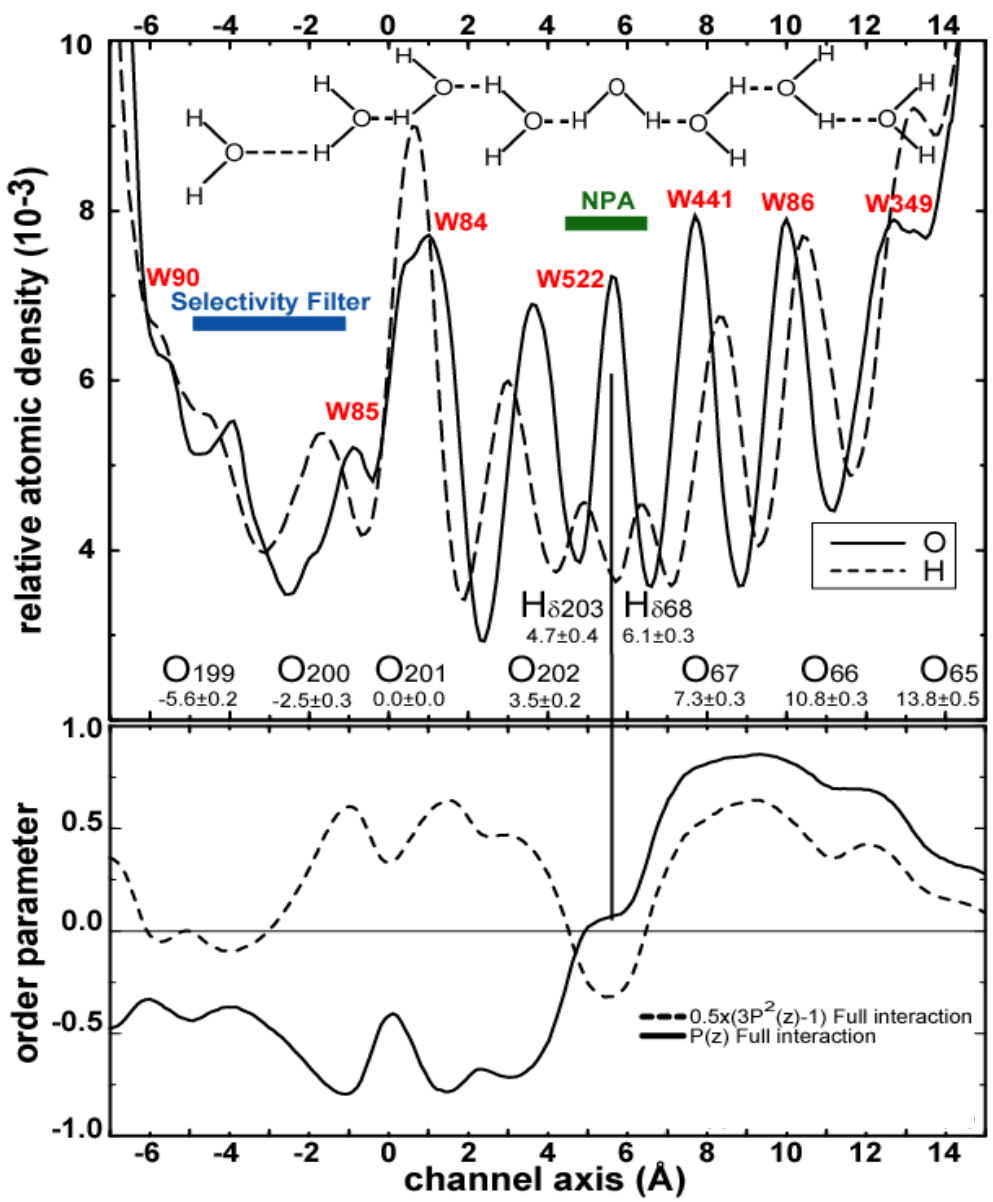
Improvement of statistics

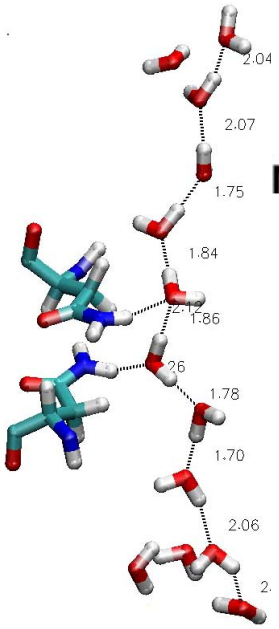
Water Bipolar Configuration in Aquaporins

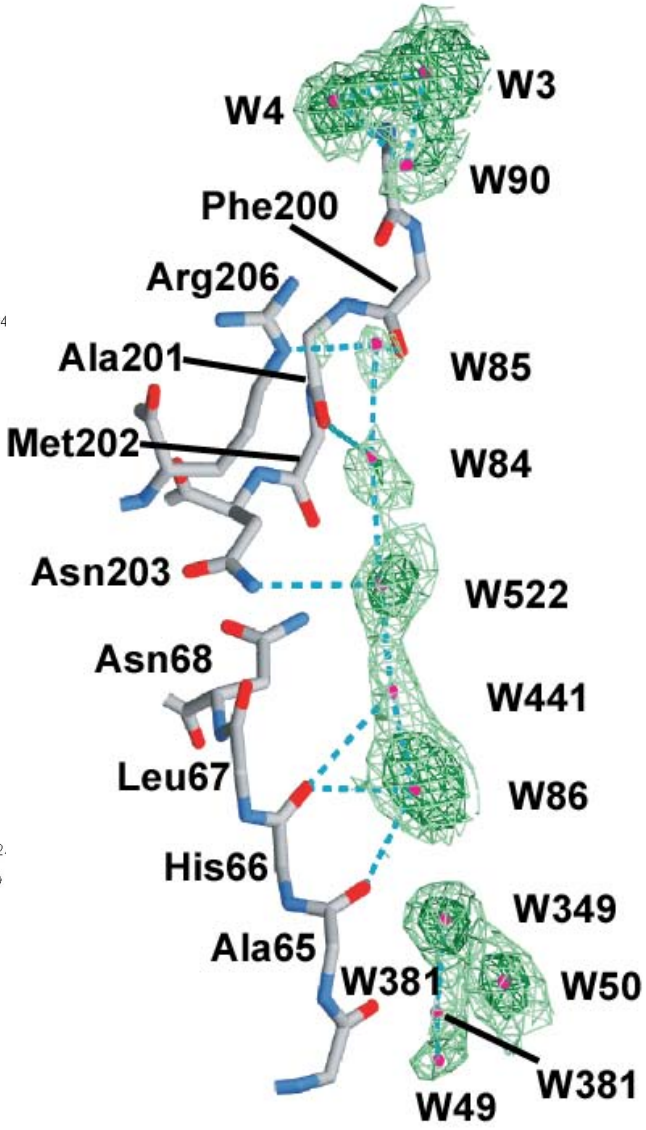


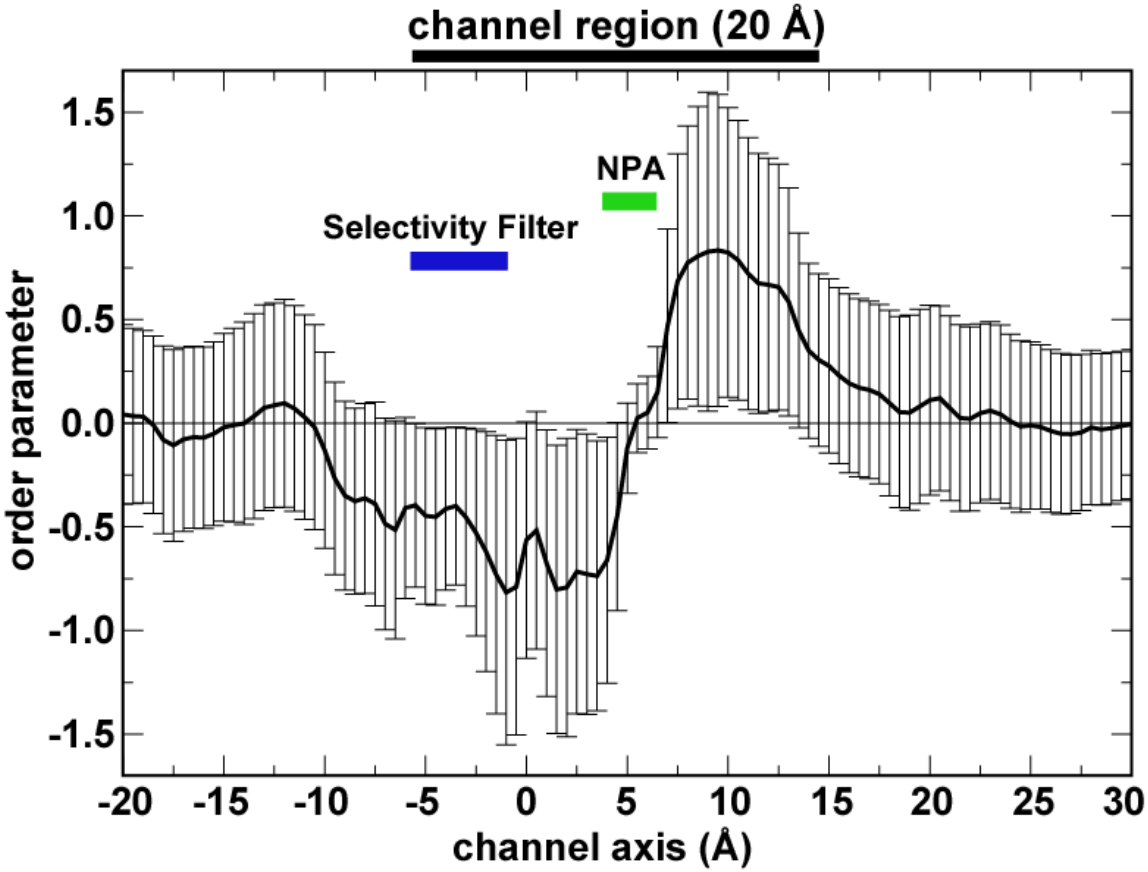


Water Bipolar Configuration in Aquaporins







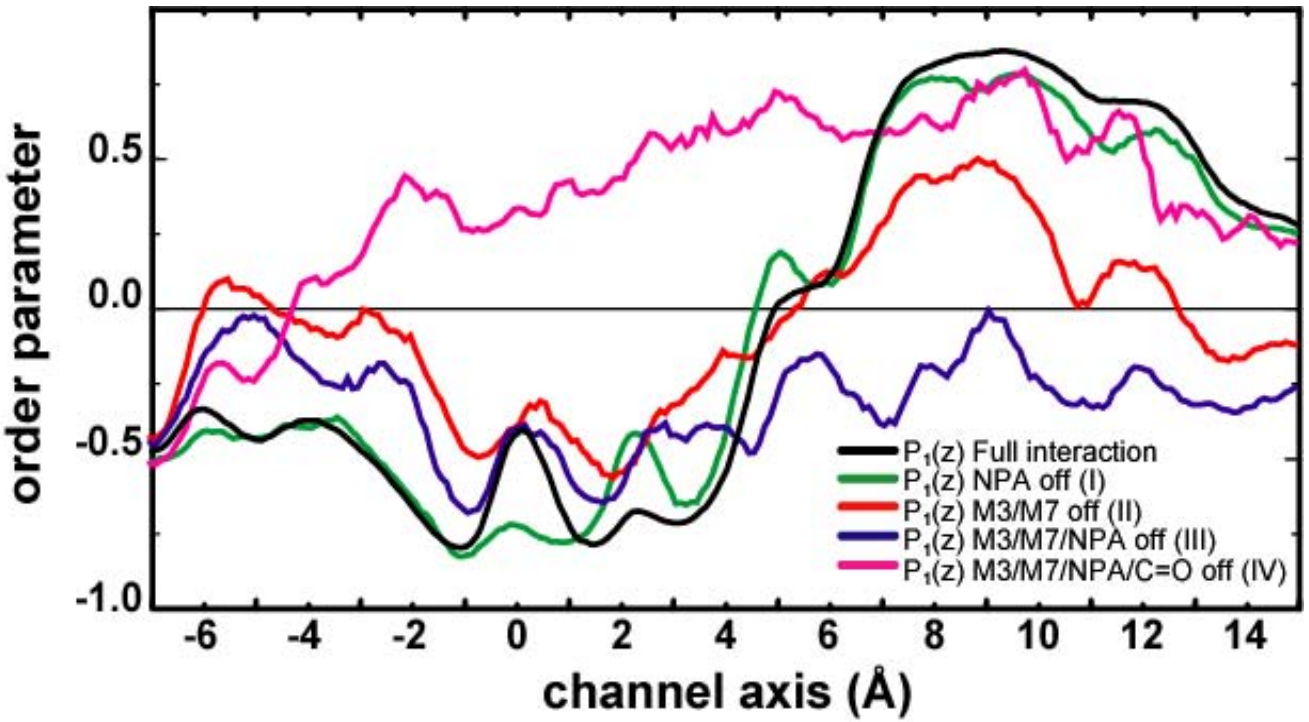

channel region (20 Å)
NPA
Selectivity Filter
order parameter
channel axis (Å)
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-20
-15
-10
-5
0
5
10
15
20
25
30

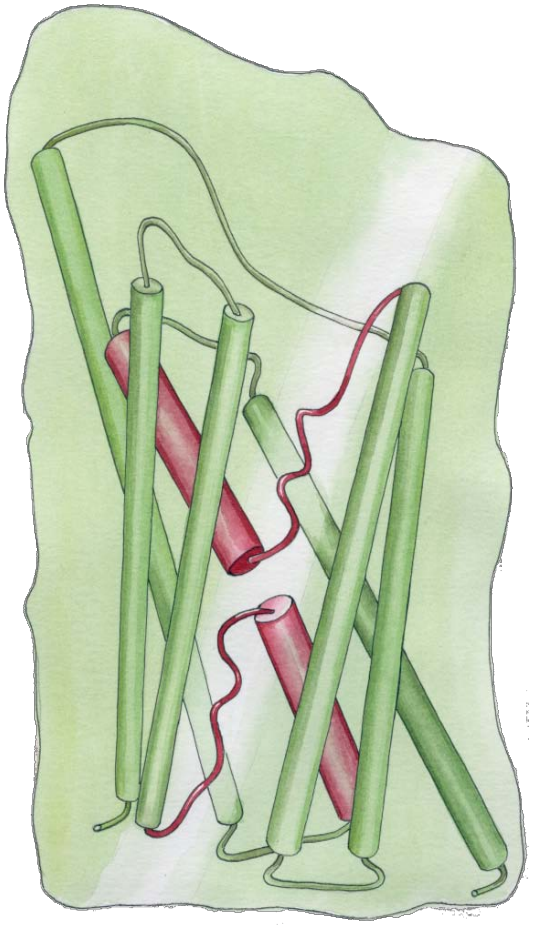
R E M E M B E R:

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!

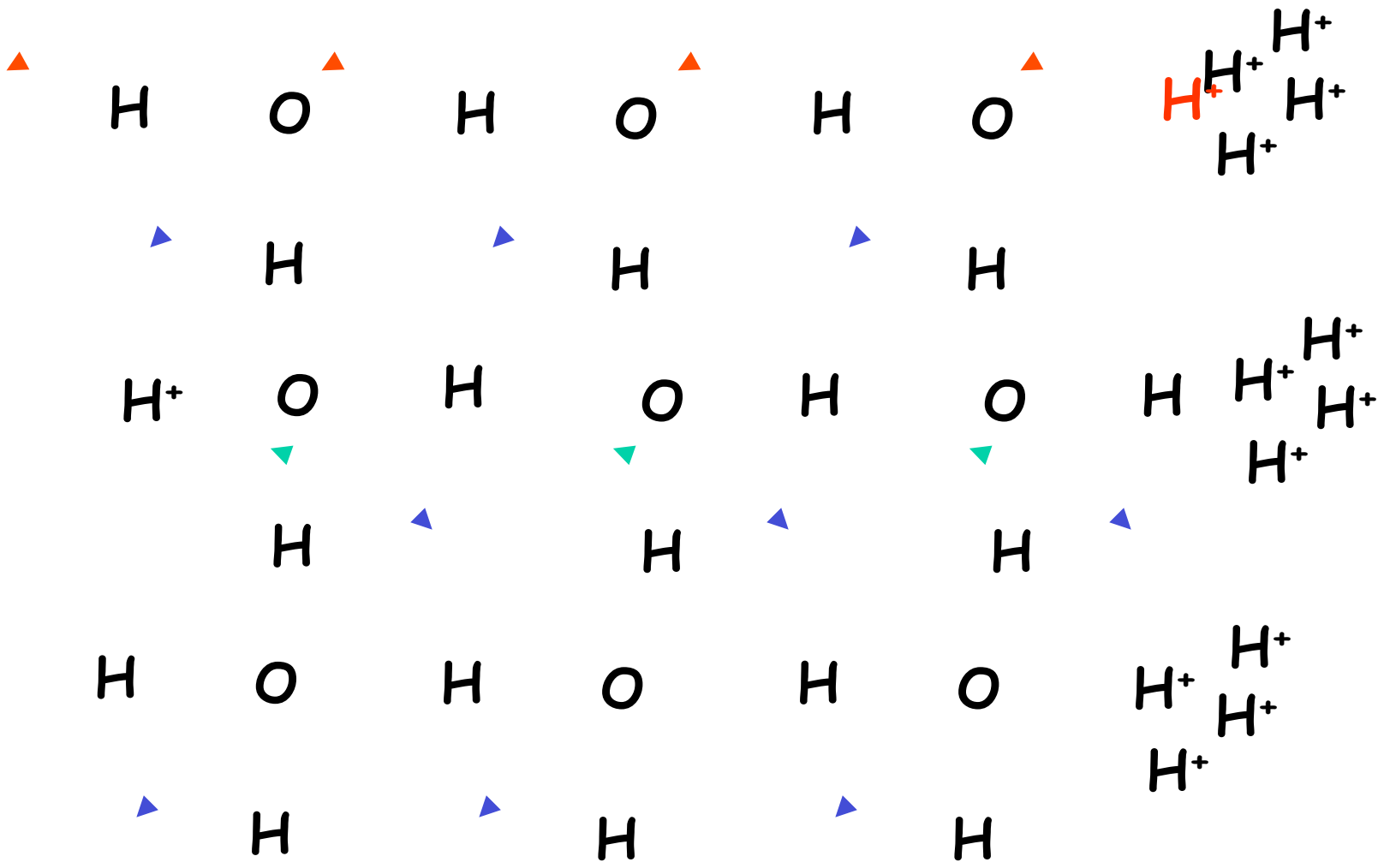
This is a powerful technique to test hypotheses developed during your simulations. Use it!

Electrostatic Stabilization of Water Bipolar Arrangement

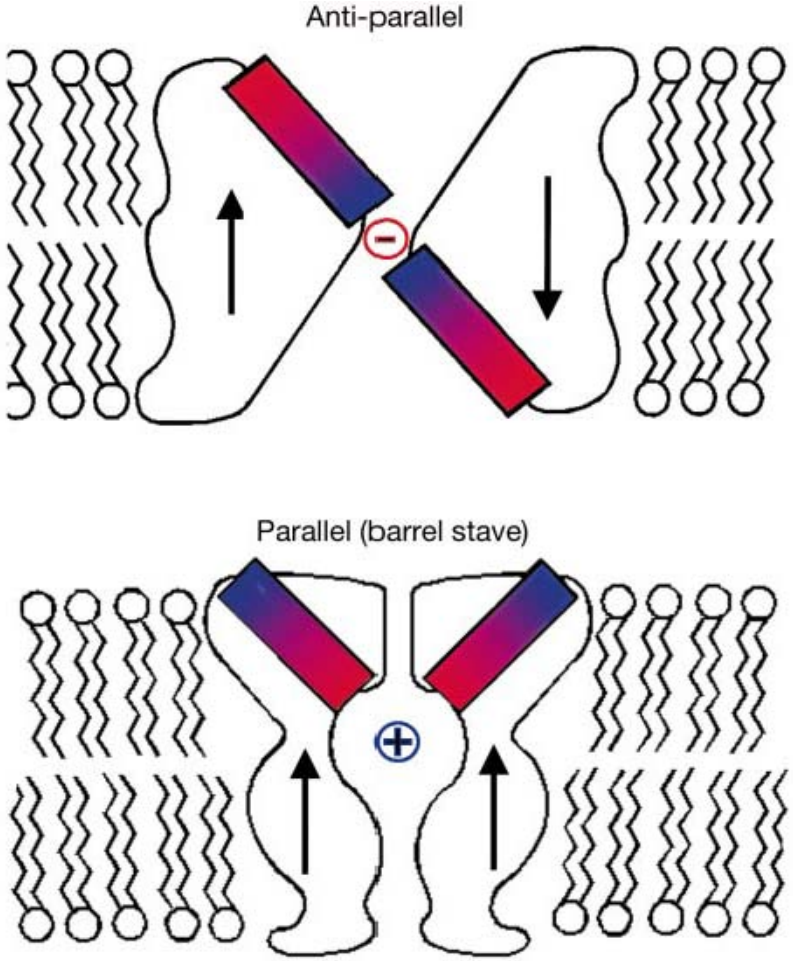




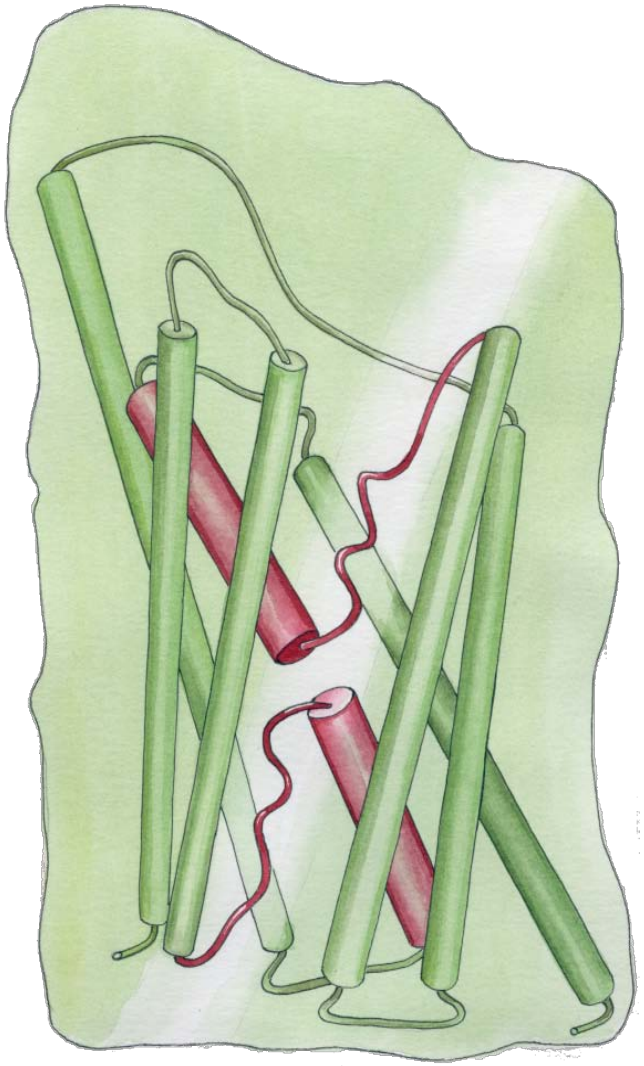
Proton transfer through water



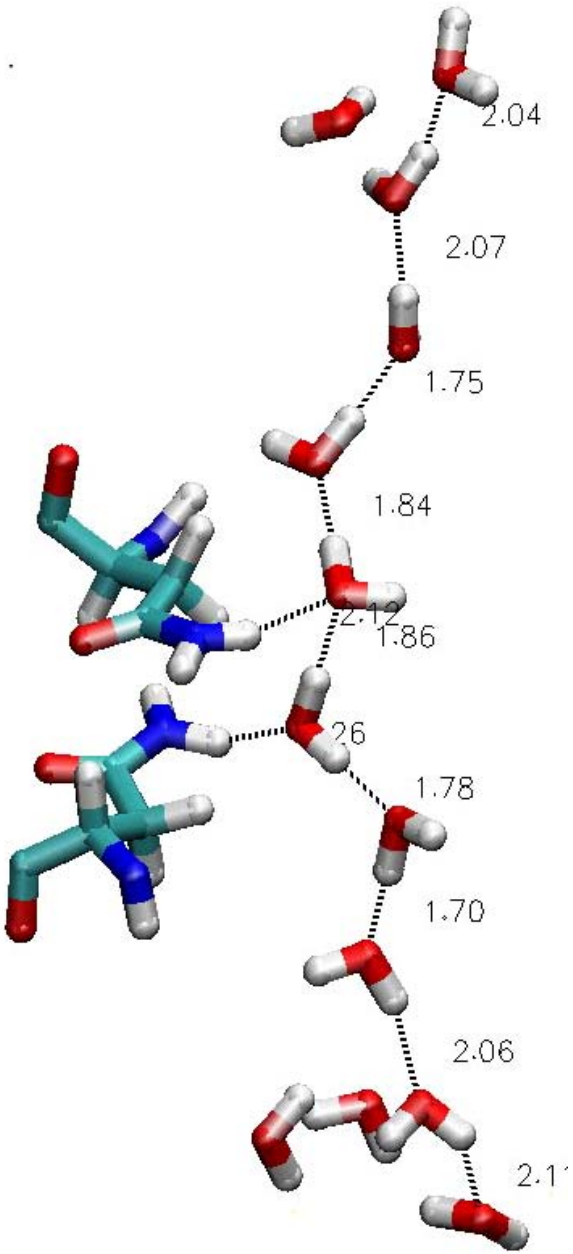
Cl^- channel

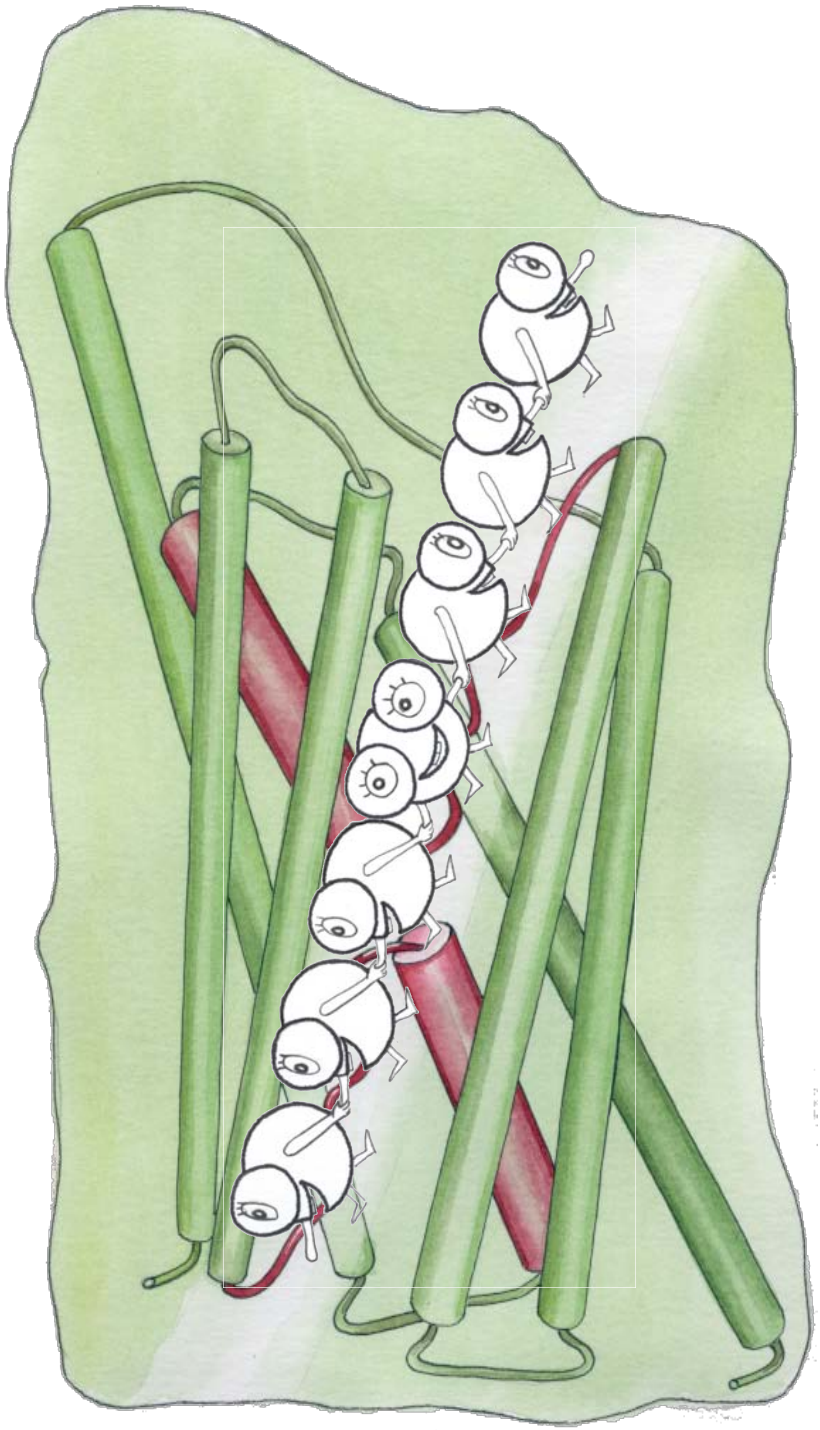


K^+ channel

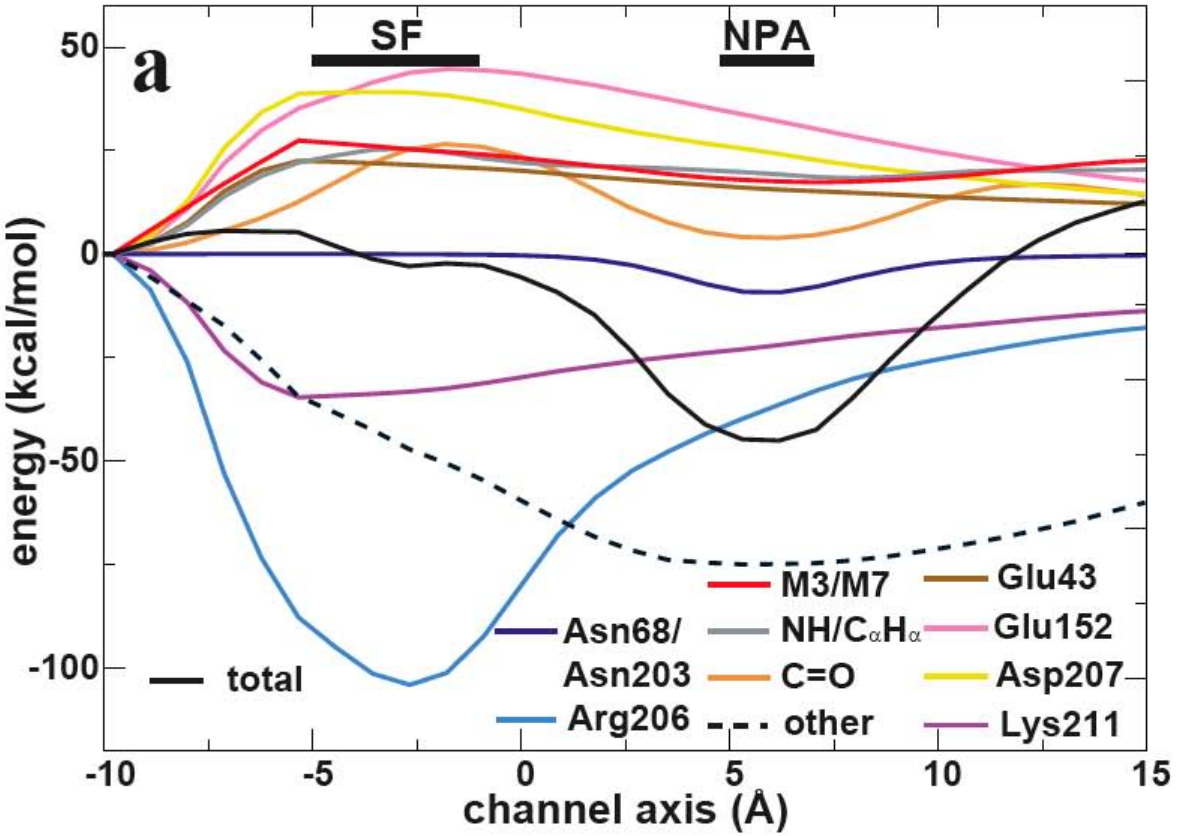


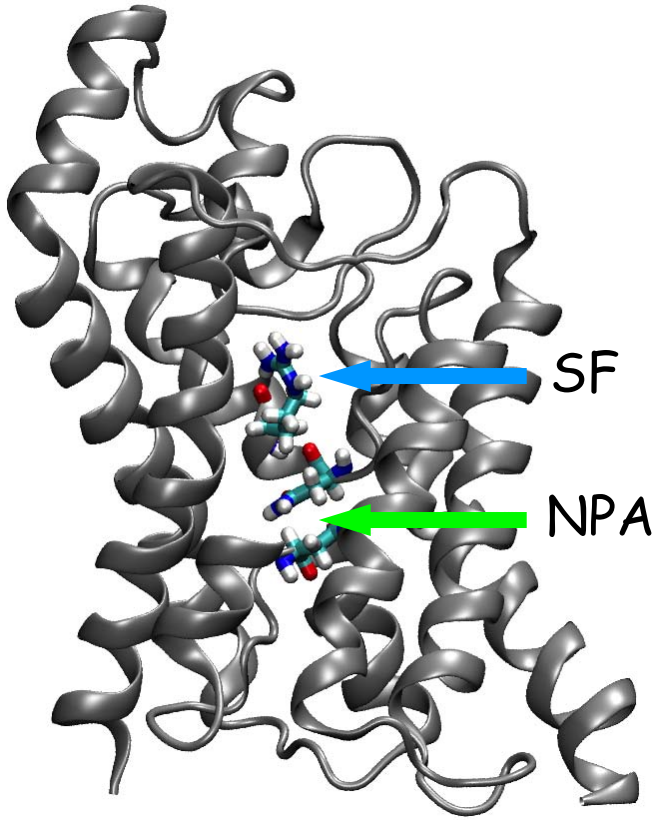
Aquaporins


2.04
2.07
1.75
1.84
2.12
1.86
26
1.78
1.70
2.06
2.1



A Complex Electrostatic Interaction

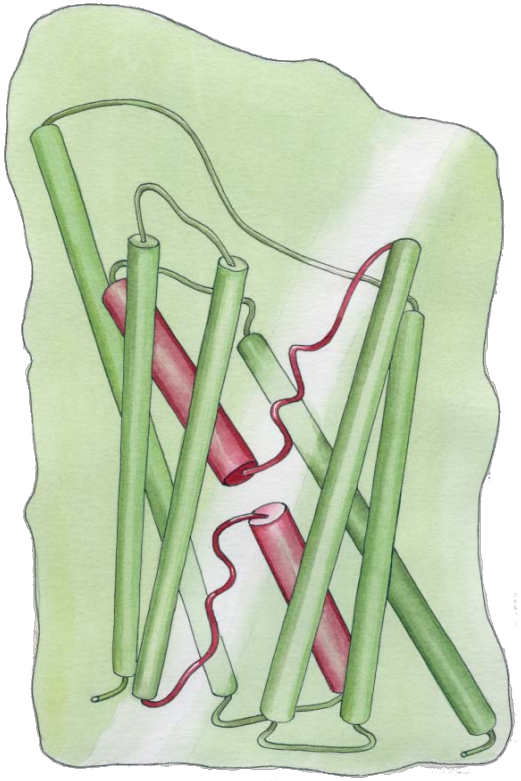


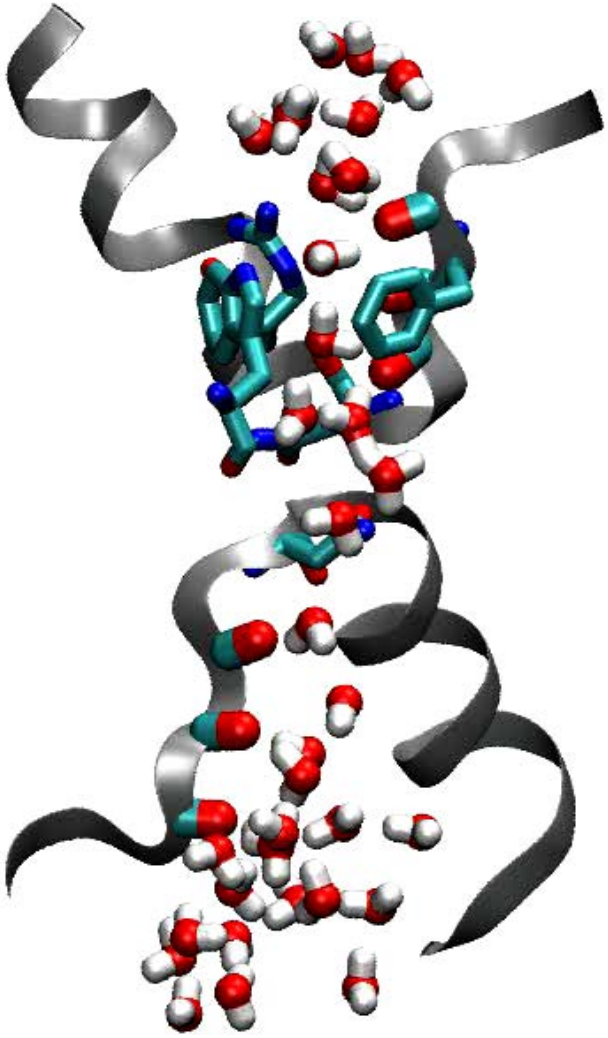


"Surprising and clearly not a hydrophobic channel"

M. Jensen, E. Tajkhorshid, K. Schulten, *Biophys. J.* **85**, 2884 (2003)

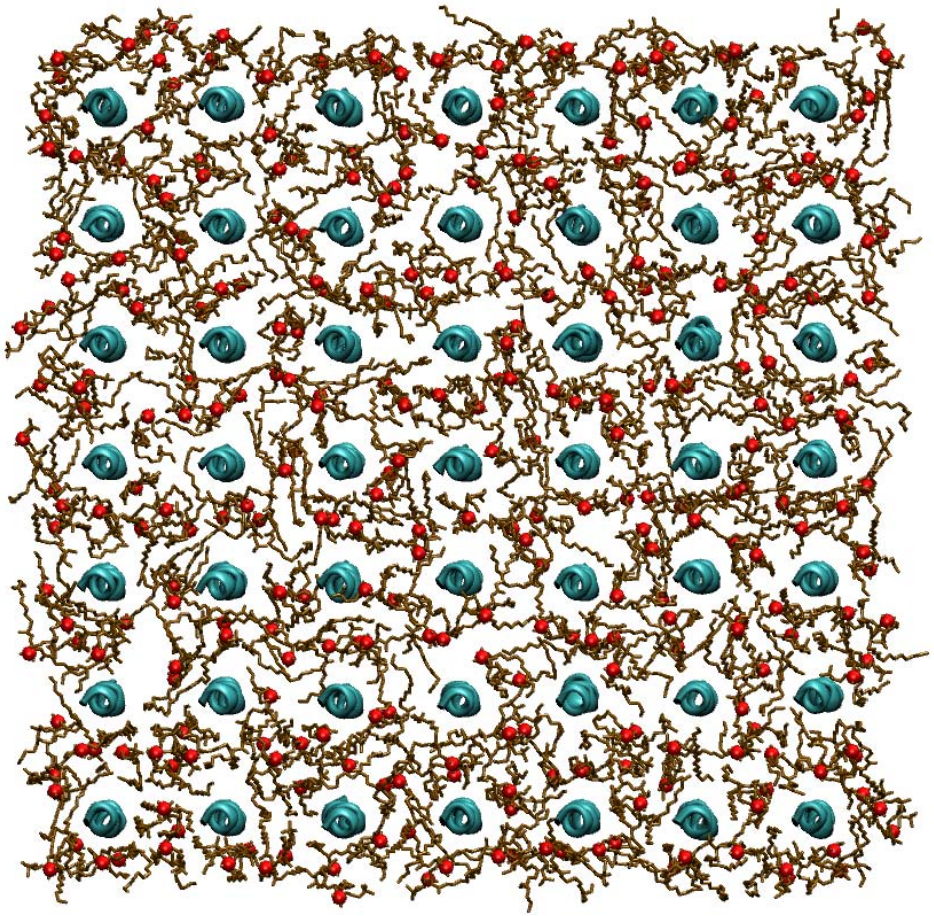
A Repulsive Electrostatic Force at the Center of the Channel



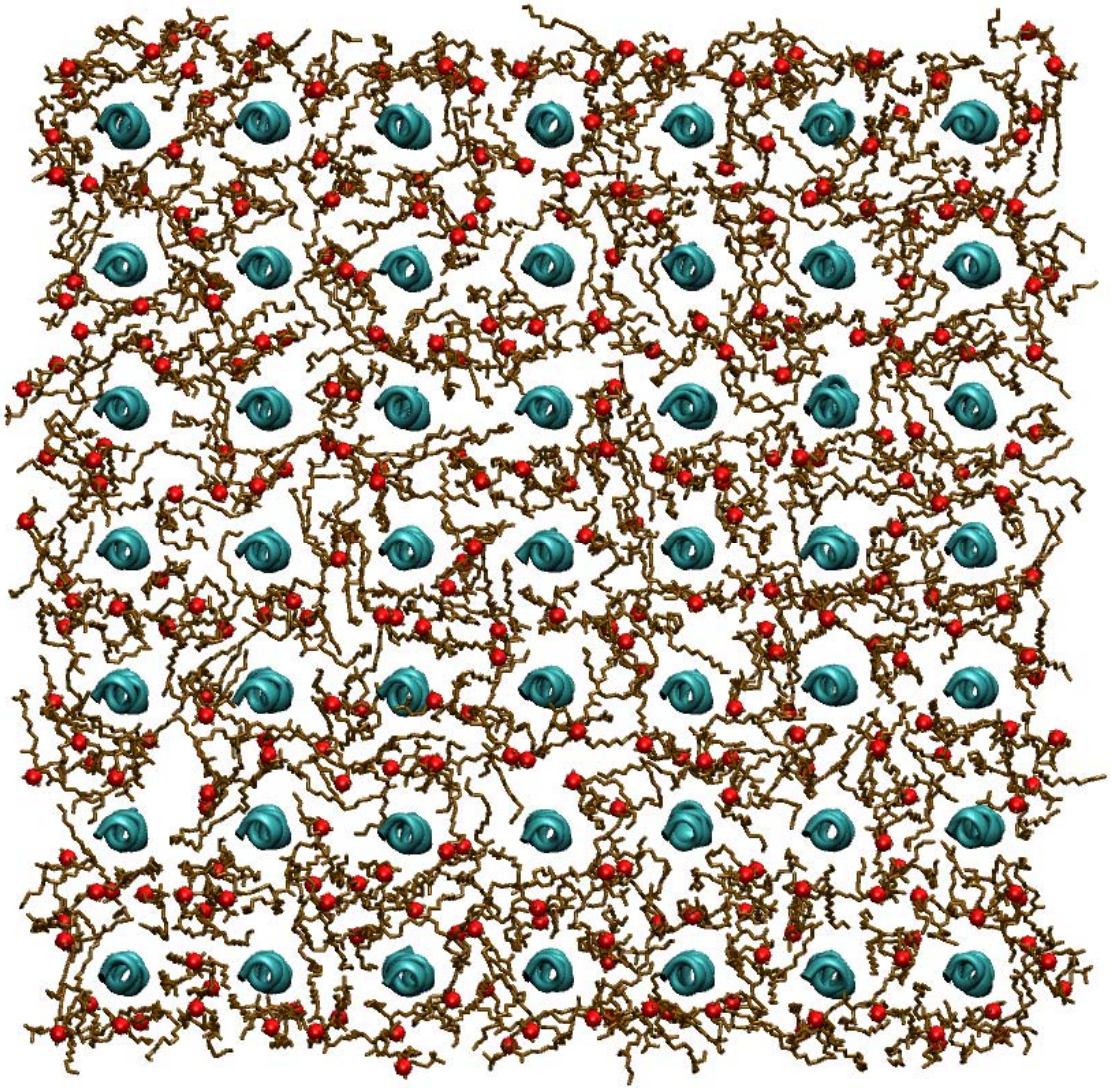


QM/MM MD of the behavior of an excessive proton

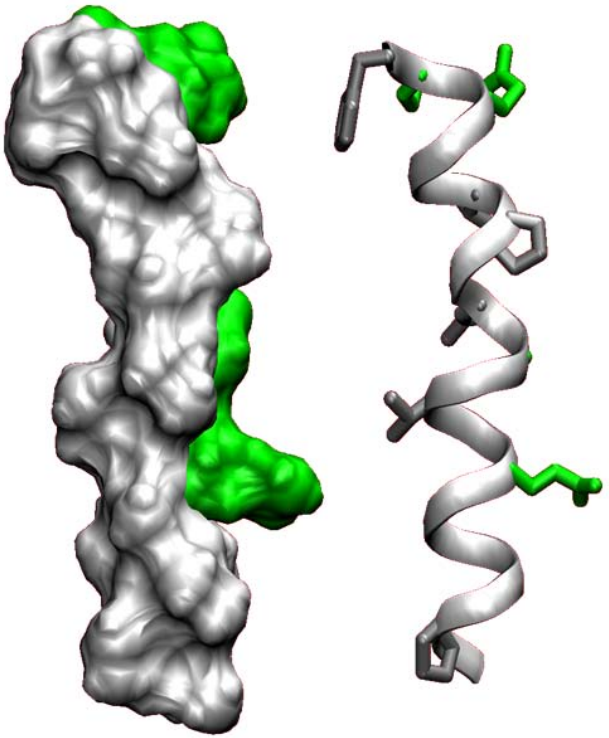
Combining all-atom and coarse-grained models to simulate transport across lipid bilayers



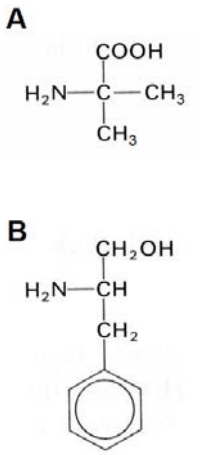
Peptide aggregation and "Pore" formation in lipid bilayers



Alamethicin



Phl20
Gln19
Glu18
Aib17
Aib16
Val15
Pro14
Aib13
Leu12
Gly11
Aib10
Val9
Aib8
Gln7
Ala6
Aib5
Ala4
Aib3
Pro2
Aib1



- 20-residue peptide
- No charge
- forming pores in the membrane

20 residue antimicrobial peptide

CG molecular systems allow for time scales of 3-4 orders of magnitude longer, because:

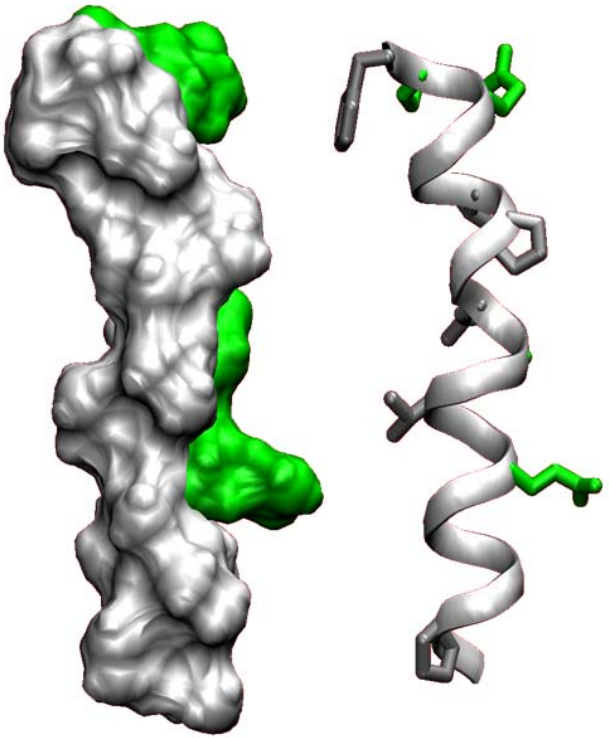
- Significant reduction of the degrees of freedom (or number of interacting particles/beads)
- Softer potentials allowing much longer time steps

μm length scale and μs time scale

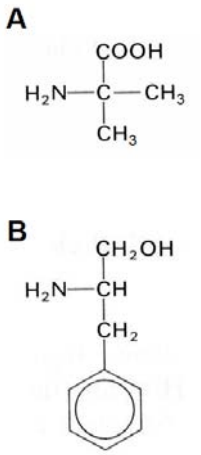
Bilayer, micelle, and vesicle formation

Fusion of bilayers and vesicles, …

Alamethicin



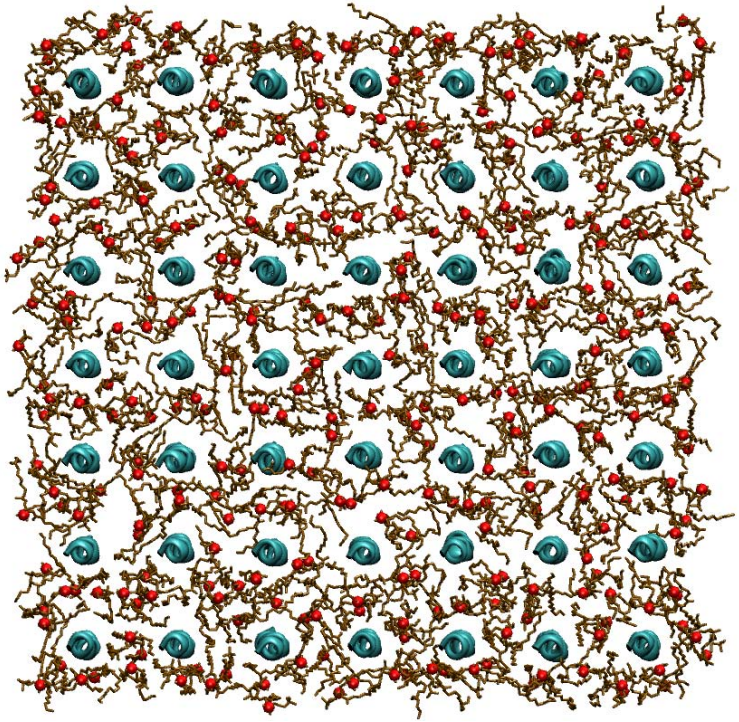
Phl20
Gln19
Glu18
Aib17
Aib16
Val15
Pro14
Aib13
Leu12
Gly11
Aib10
Val9
Aib8
Gln7
Ala6
Aib5
Ala4
Aib3
Pro2
Aib1

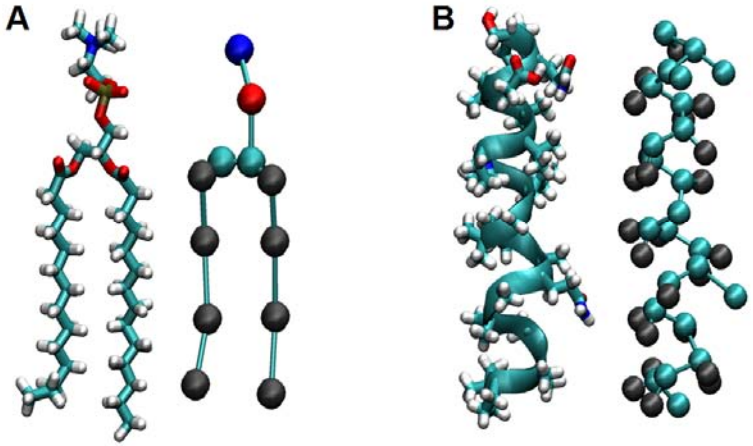


- 20-residue peptide
- No charge
- forming pores in the membrane

20 residue antimicrobial peptide

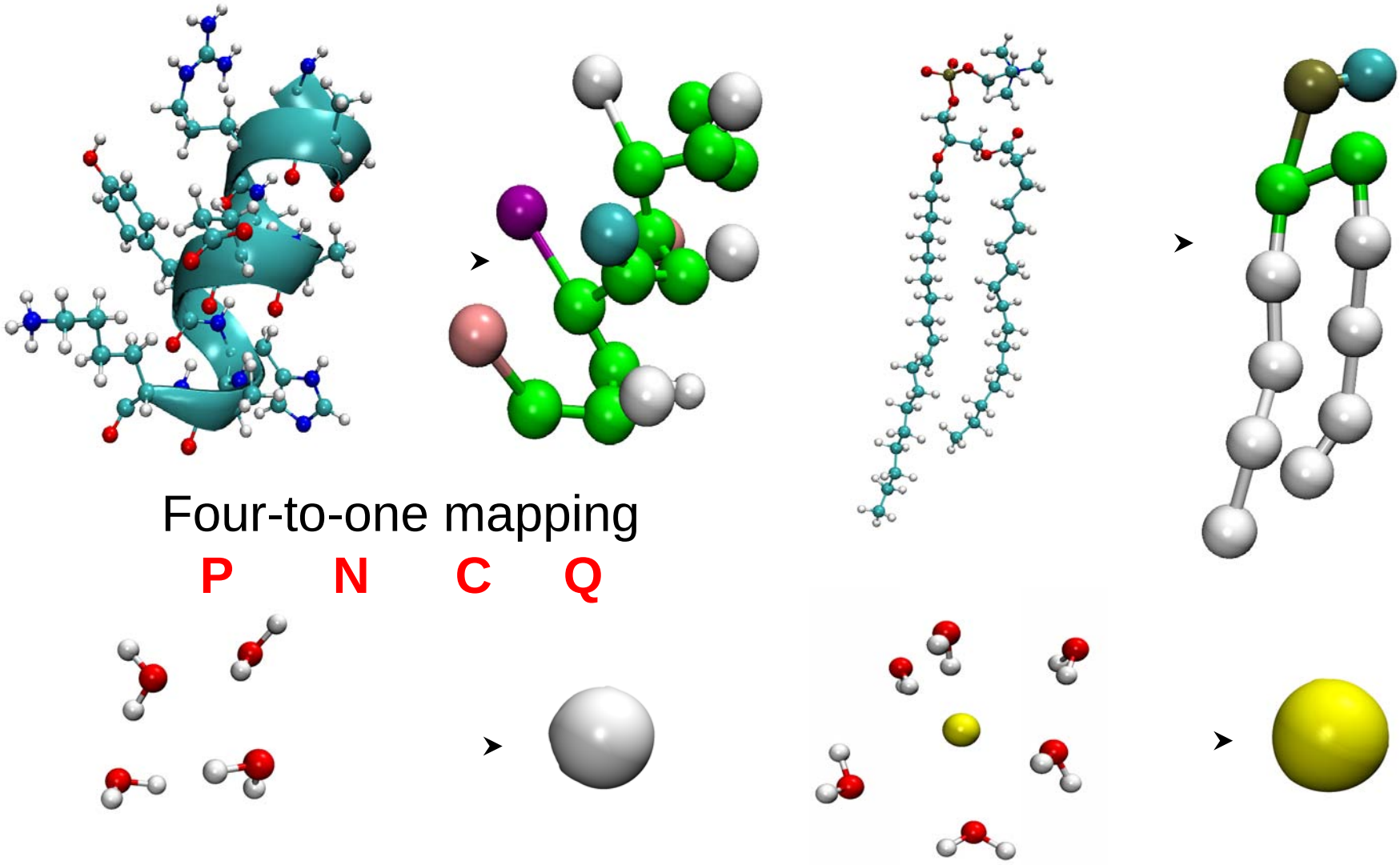
Simulation Setting





- 49 peptides in 288 DMPC
- All-atom model equilibrated 1ns
- Converted to a CG model
- Simulated for 1μs
- 0.5 μs snapshot was reverse-corarse-grained to an all atom model
- All atom model simulated for 20 ns

Coarse-Graining

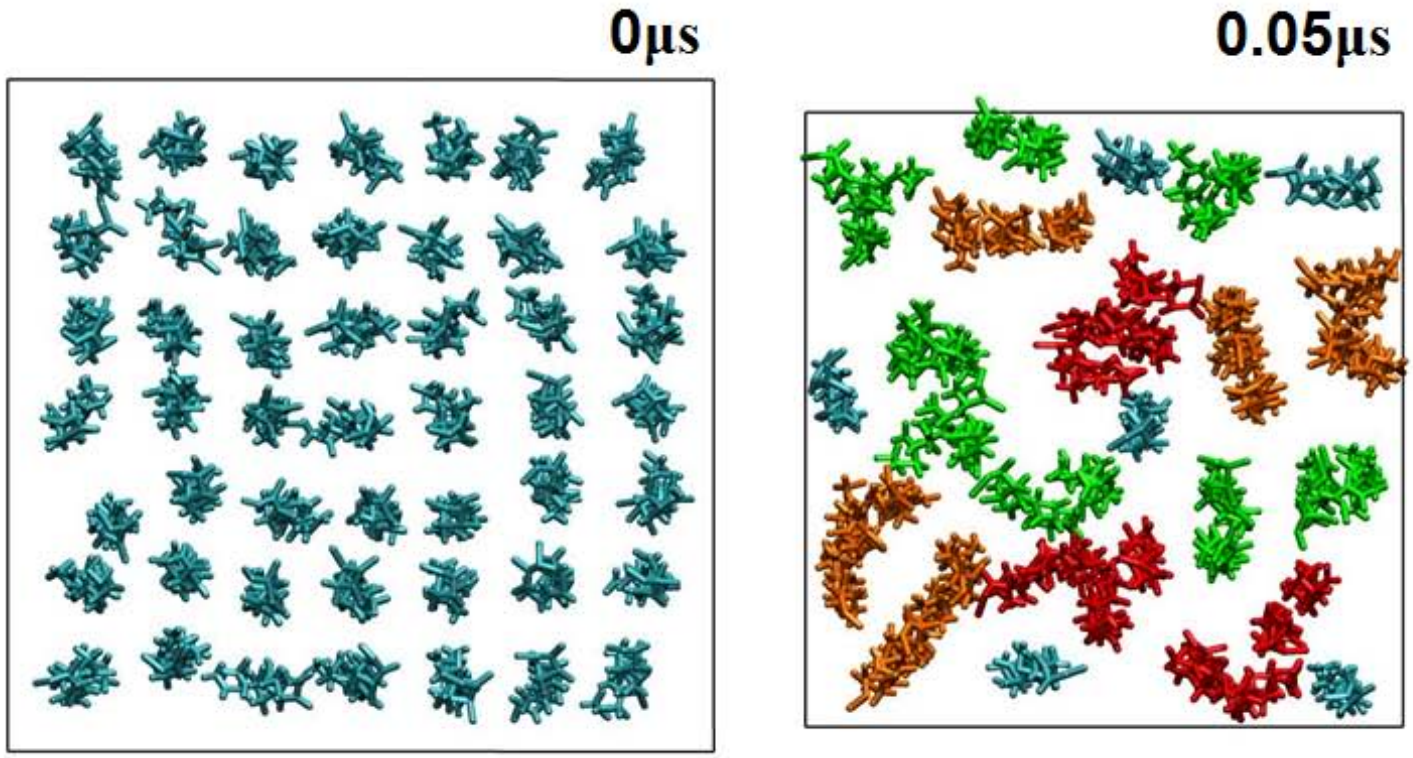


Four-to-one mapping

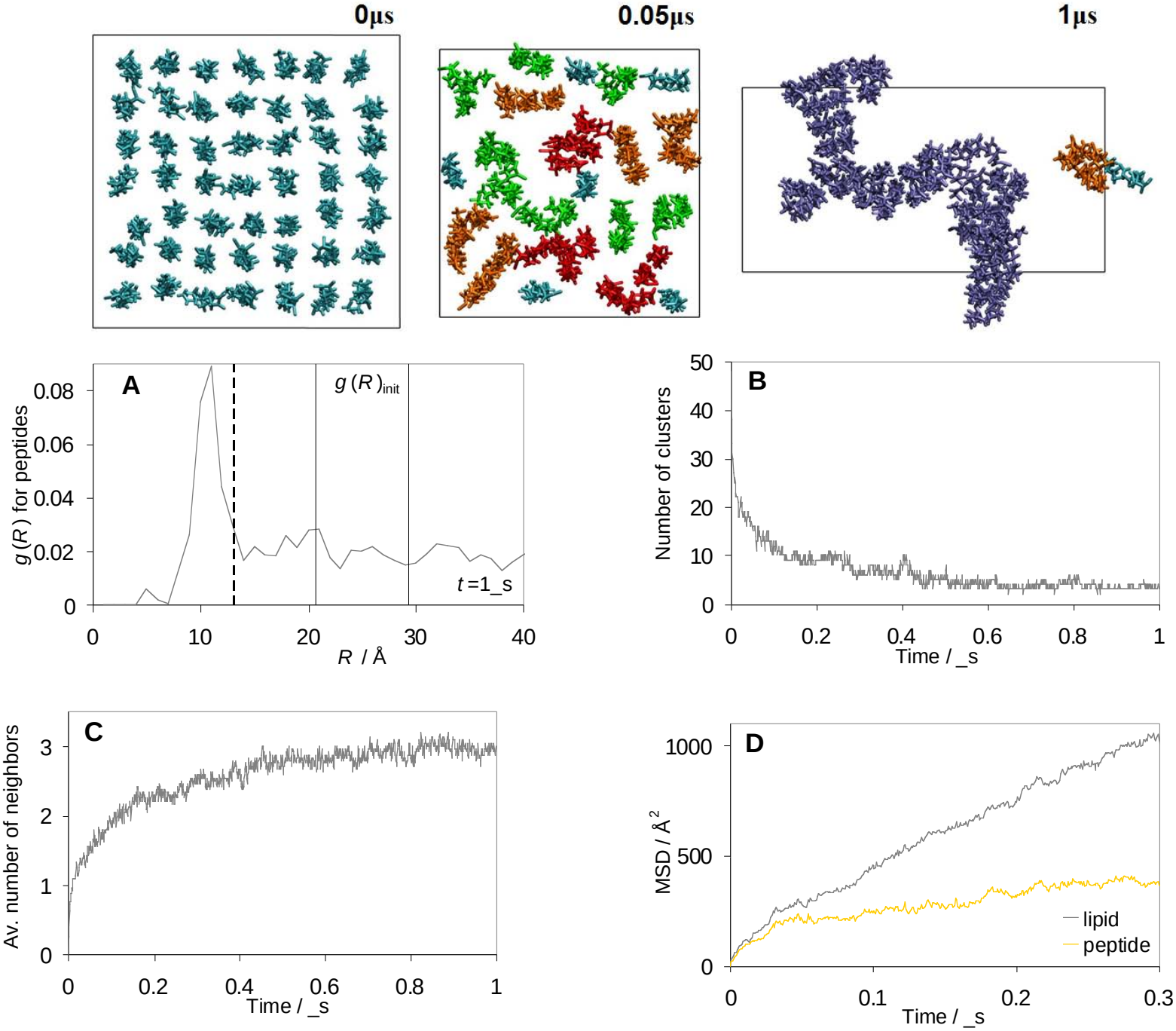
P N C Q

Employs CHARMM-like force fields
parameterized using all-atom simulations

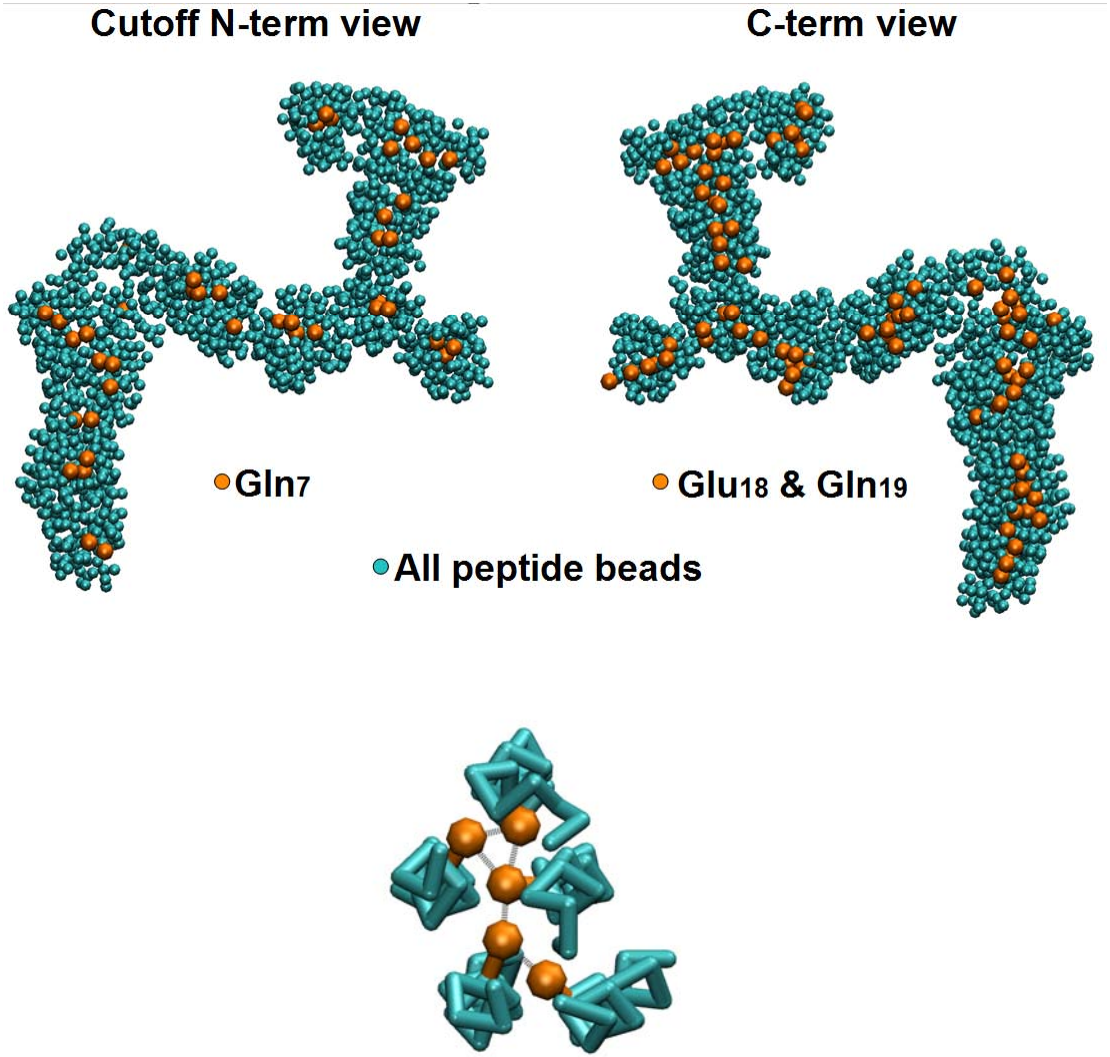
Peptide Aggregation

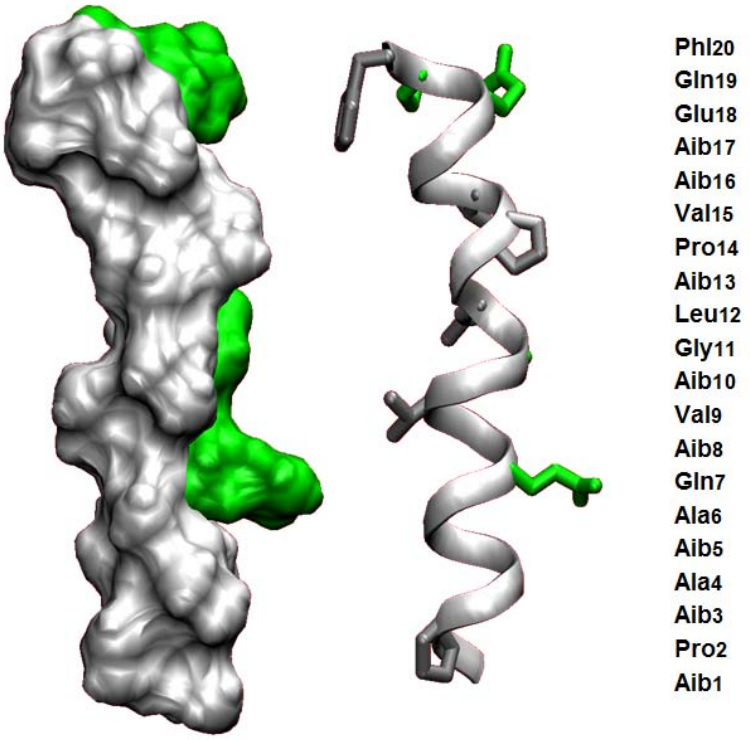


L. Thogersen, et al., ***Biophysical Journal (***2008)

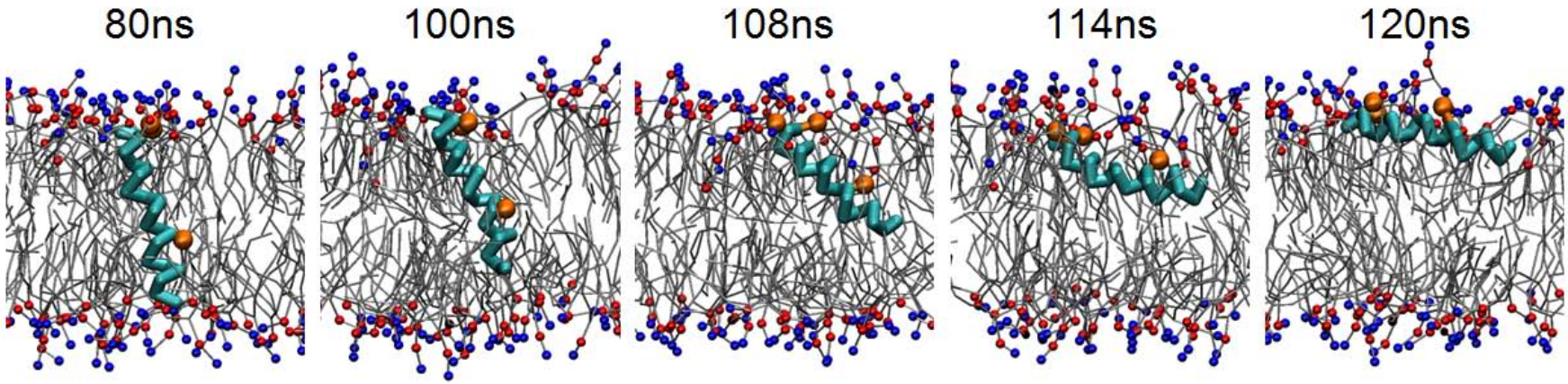

0μs
0.05μs
1μs
A
g (R)init
t =1_s
g(R) for peptides
R / Å
B
Number of clusters
Time / _s
C
Av. number of neighbors
Time / _s
D
MSD / Å2
lipid
peptide
Time / _s

What brings the helices together?



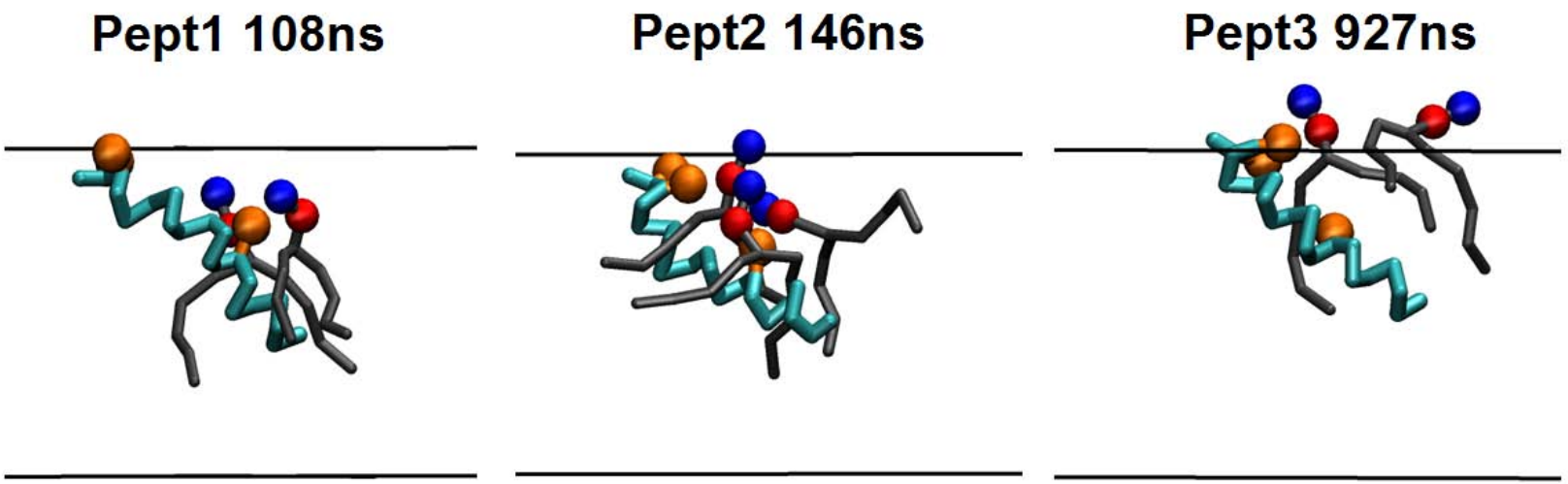


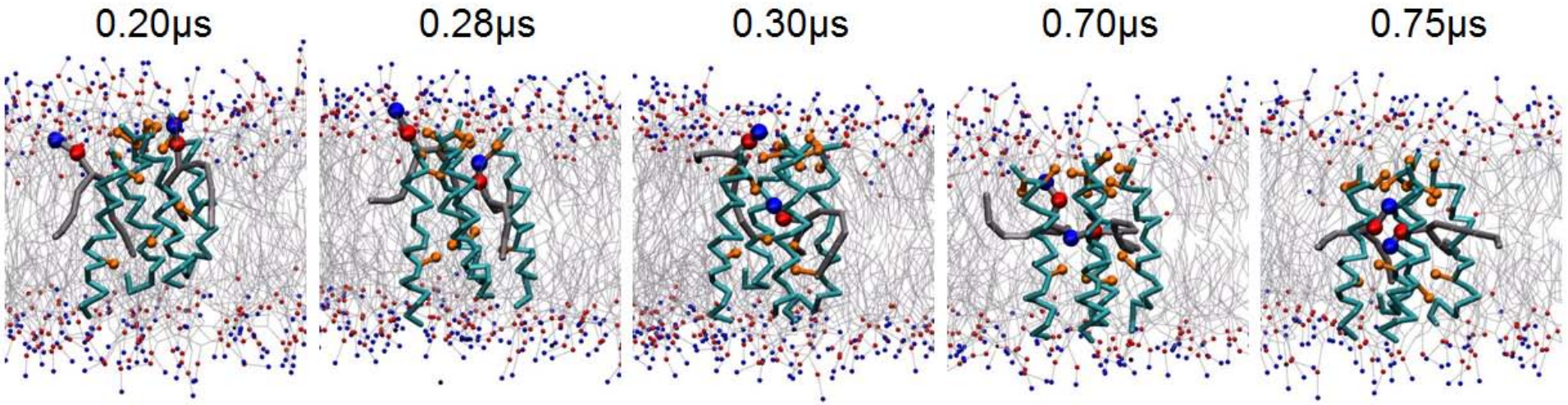
Peptide Insertion



L. Thogersen, et al., ***Biophysical Journal (***2008)

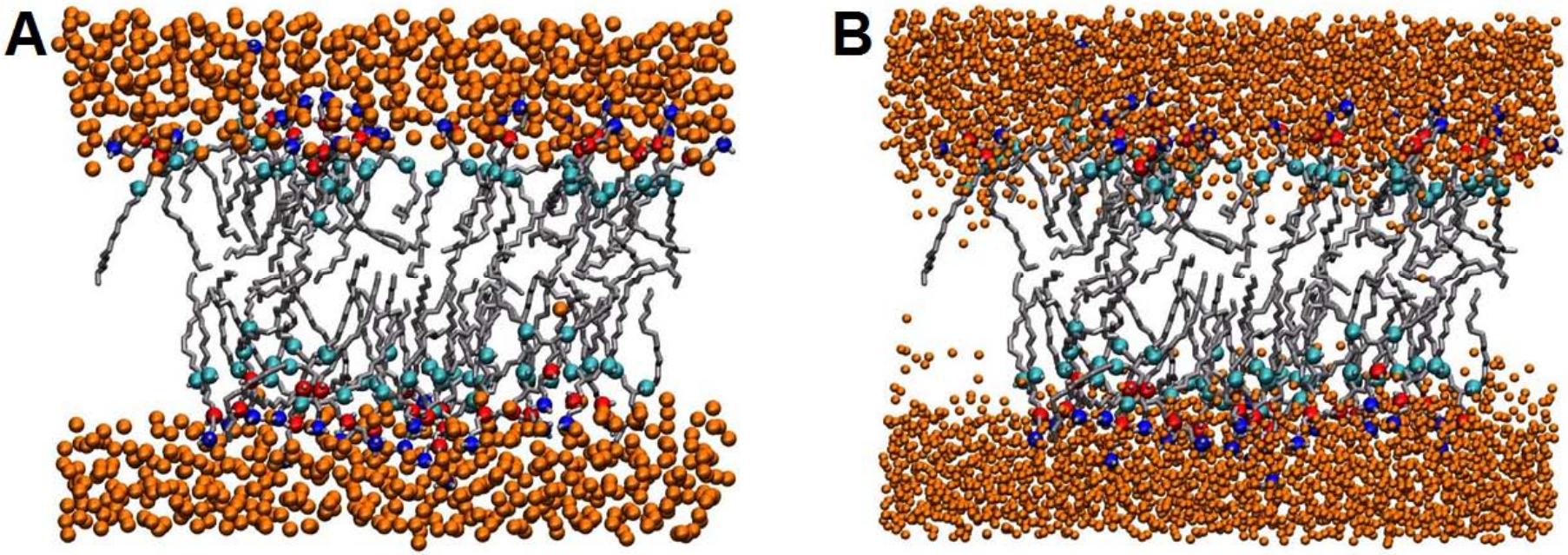
Strong perturbation of the bilayer structure



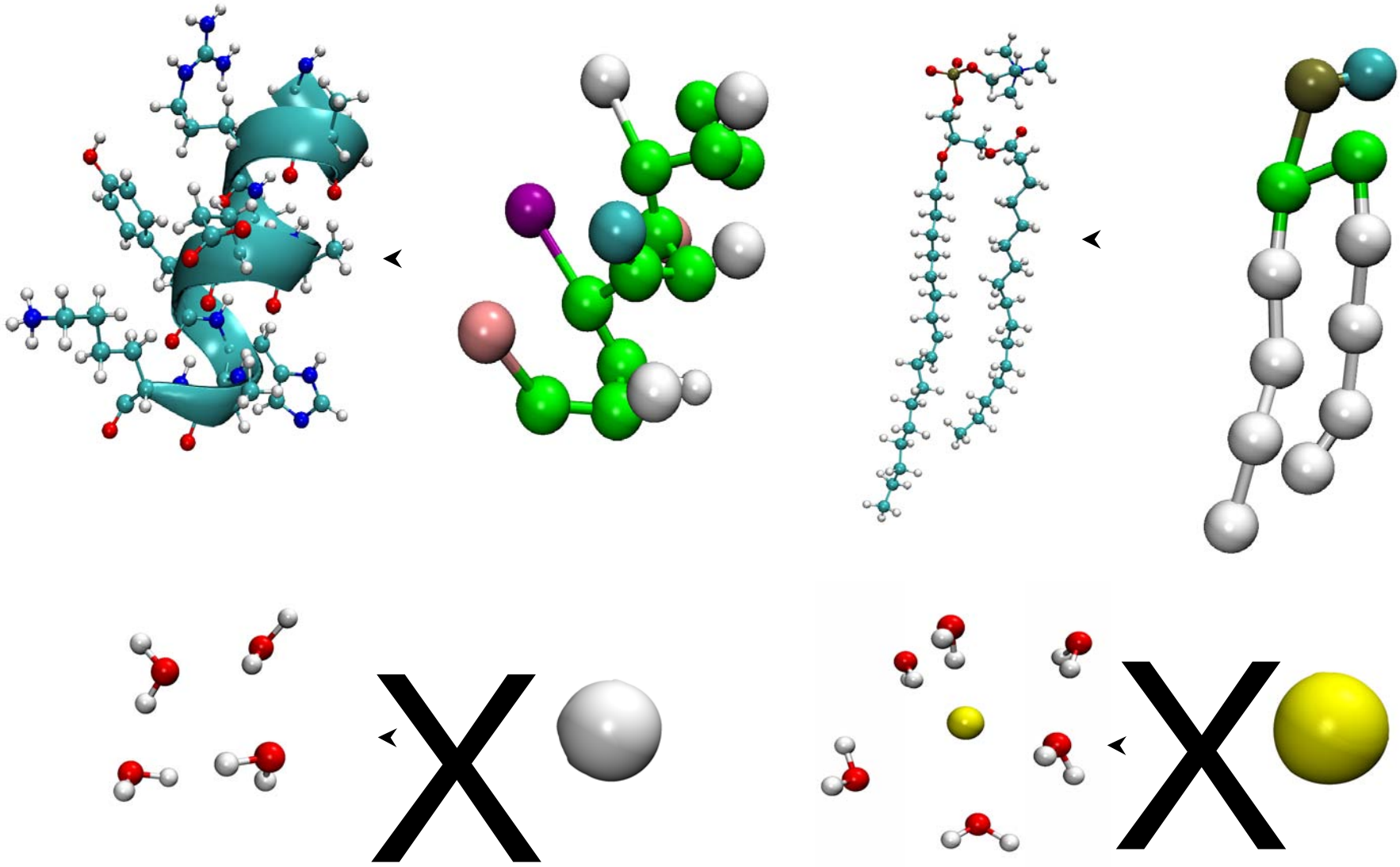


BUT no water permeation/pore formation!?

Hydration of the head group region in coarse-grained and all-atom models



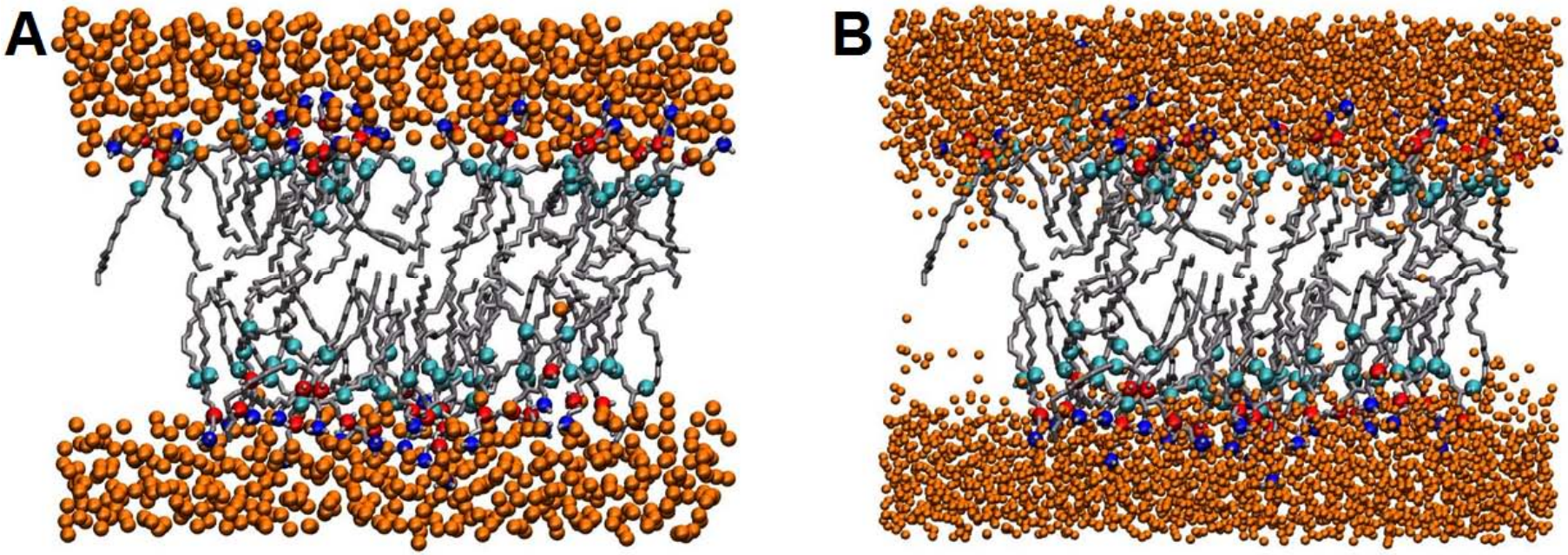
Reverse Coarse-Graining



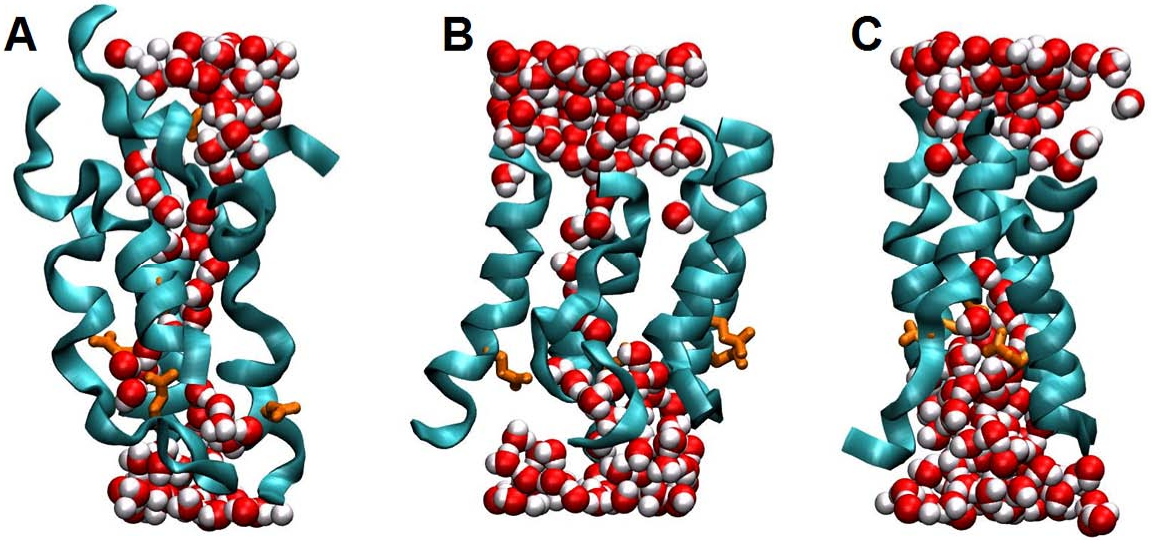
Reverse Coarse-Graining

- Mapping back CG beads to all-atom clusters
- *Re-solvating the system*
- 5000 steps of minimization
- Simulated annealing for 20 ps (T changing from 610K to 300K, $\Delta T = -10K$) while constraining atoms to the position of the corresponding CG beads

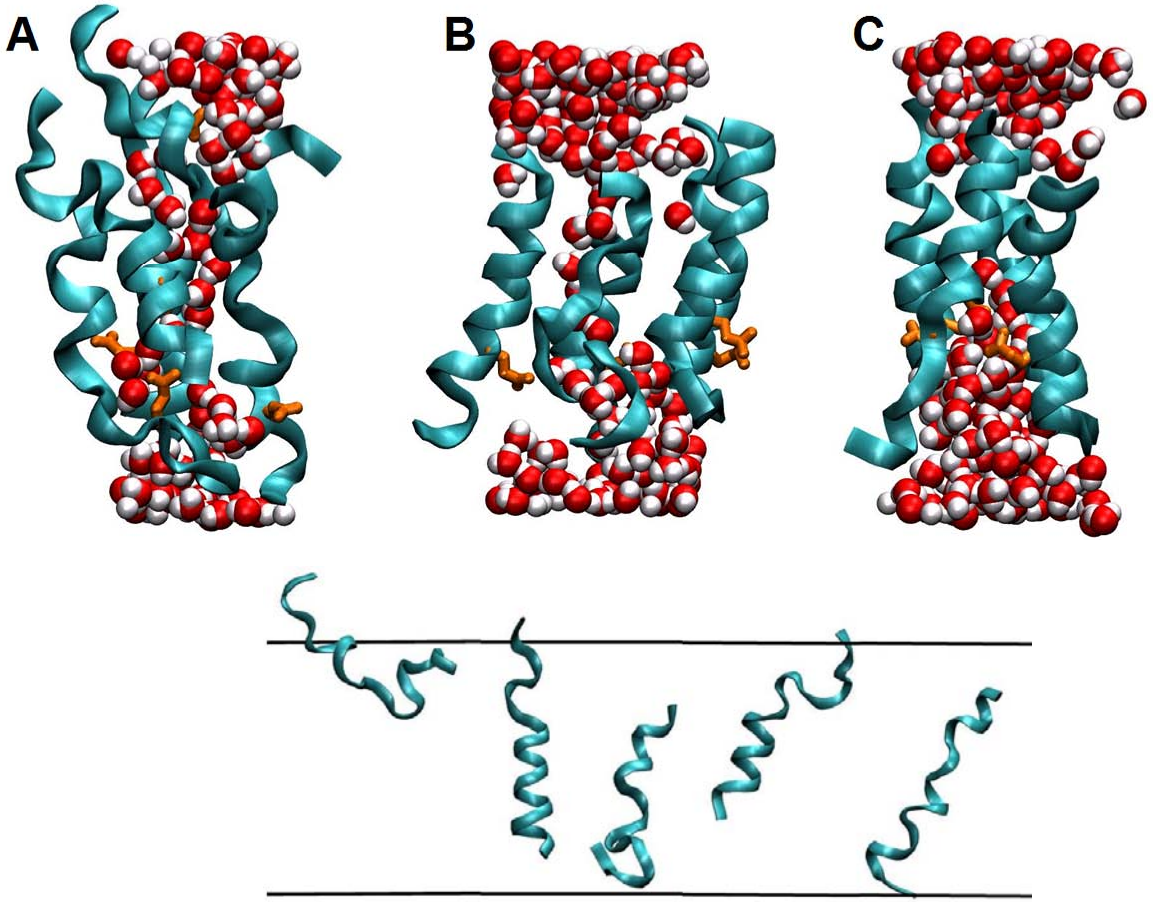
Hydration of the head group region in coarse-grained and all-atom models



Alemethicin-induced Membrane Poration



Alemethicin-induced Membrane Poration



Alemethicin-induced Membrane Poration