

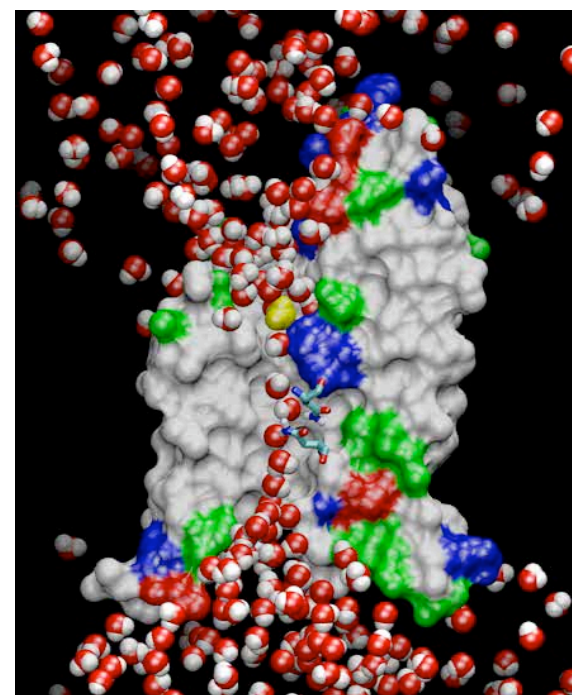
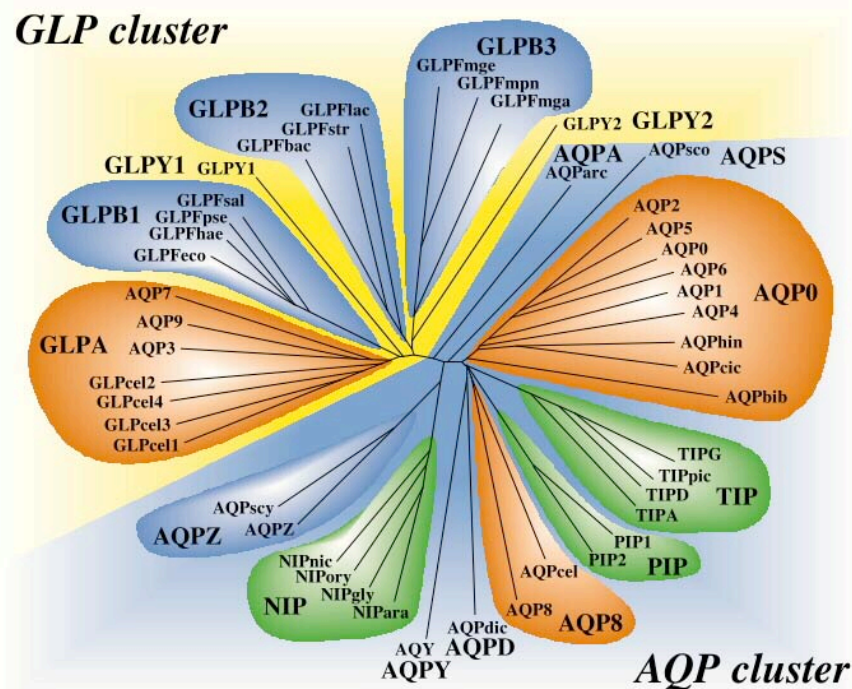
Physical Bioinformatics - A Case Study

Sequence and structure information are the bedrock on which an understanding of cellular functions and the underlying physical mechanisms can be built. This lecture illustrates how the two sources of information are combined to investigate by means of the program VMD function and mechanism of the aquaporin family of membrane channels that transport water and certain small solutes across cell walls. Introducing first the key architectural features of a single aquaporin, structures and sequences of four aquaporins are aligned and common features recognized. The shared and distinct features are examined closely and used as guideposts leading quickly to key questions regarding the mechanism underlying aquaporin's efficient conduction and selection. The questions are addressed by means of molecular dynamics simulations using the program NAMD that reveal the physical principles behind water transport and highly selective solute co-transport in aquaporins. Sequence-structure information is viewed again to elucidate tetramer binding and pathologies connected with certain aquaporin mutants. The lecture introduces the concepts behind the programs employed and emphasizes those aspects of the case study that can be applied for investigations of other protein families.

Physical Bioinformatics - A Case Study

Aquaporin Family of Membrane Channels

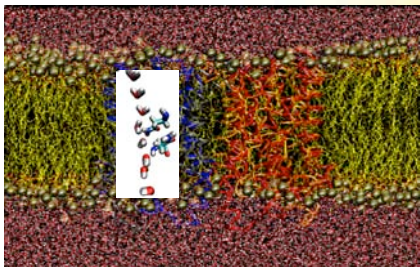
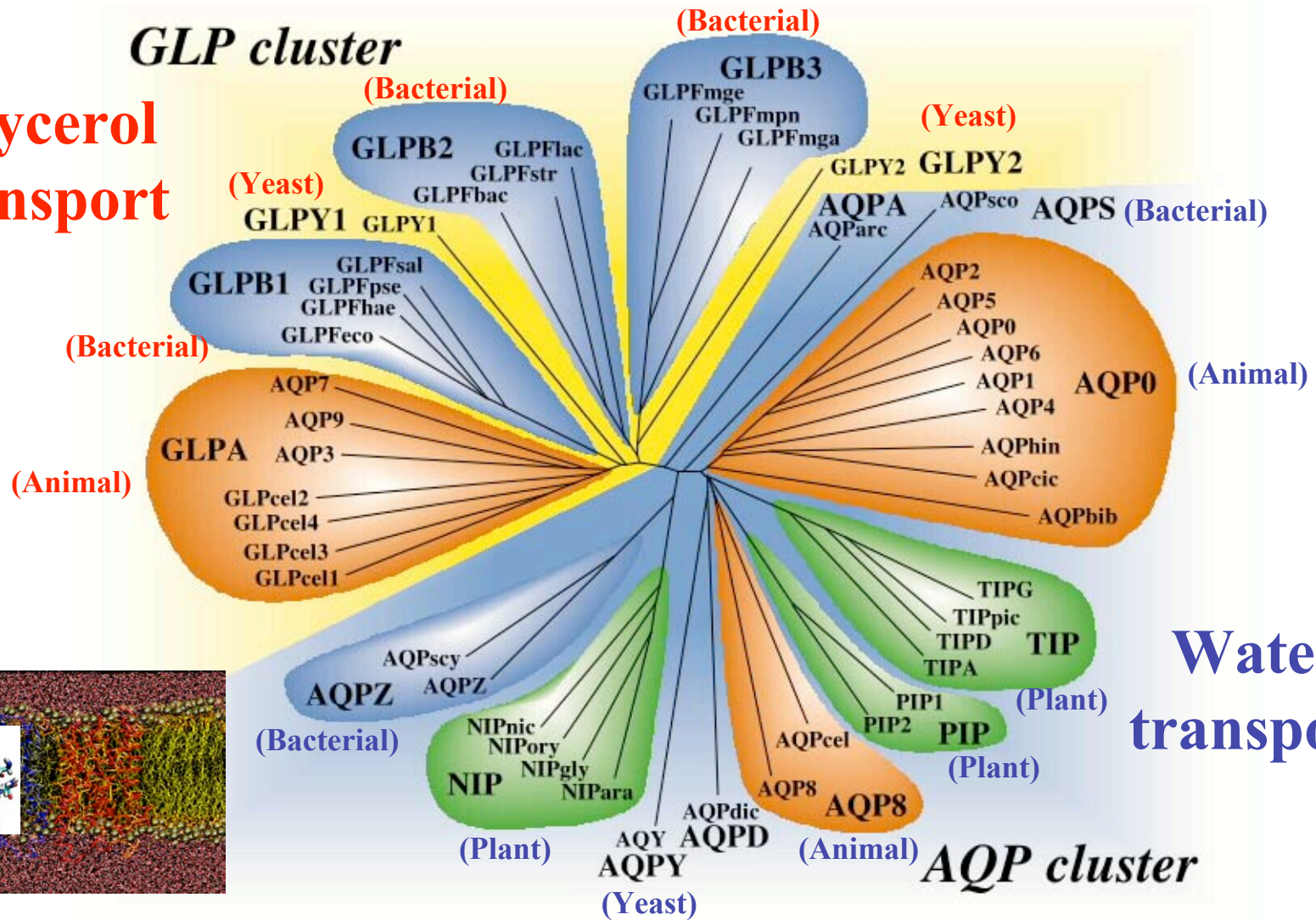
Klaus Schulten, U. Illinois at Urbana-Champaign



AQP0_HUMAN	---	LNTLHPAVSV	GQATTVEIF	LT	LQFVLCIFATYDE	-RRNGQLGSVALAVGFS	LALGHLFGMY	YT	GAGM	183					
AQP1_HUMAN	---	RNDLADGVNS	GQGLGIEII	IGTLQ	LVLCVLATTDR	-RRRDLGGSAPLAIGL	SVALGHLLAIDY	TGCGI		191					
AQP2_HUMAN	---	VNALSNST	TAGQAVTVEL	FLT	LQLVLCIFASTDE	-RRGENPGTPALSIGF	SVALGHLLGIHY	TGCSM		183					
AQP3_HUMAN	G	IFATYPSG	HLDMINGFFDQ	FIGTASL	IVCVLAIVD	PYNNPVPRGLEAFTV	GLVVLVIGT	SMGFNS	GYAV	214					
AQP4_HUMAN	---	VTMVHGNL	TAGHGLLVEL	IITFQ	LVFTIFASCDS	-KRTDVTGSIALAIGF	SVAIGHLFAINY	TGASM		212					
AQP5_HUMAN	---	VNALNNNT	TQGGAMVVEL	ILTFQ	LALCIFASTDS	-RRTSPVGS	PALSIGLSVT	L	GHLVGIYFTGCSM	184					
AQP6_HUMAN	---	INVVRNSV	STGQAVAVE	LLLTQ	LVLCVFASTDS	-RQTS--GSPATMIGI	SWALGHLIGIL	FTGCSM		195					
AQP7_HUMAN	G	IFATYLP	DHMTLWRG	FLNEAWLT	GMLQLCLFAITD	QENNPALPGTEALV	IGILVVIIGV	SLGMNTGYAI		225					
AQP8_HUMAN	-	AAFVTVQEQ	GVAGALVAE	II	LTLLALAVCMGAIN	--EKTKGPLAPFSIG	FAVTVDILAGGP	VS	GGCM	209					
AQP9_HUMAN	H	IFATYPAPY	LSLANAFADQ	VVATMILLI	IVFAIFDSRNLGAPRGLEPIA	IGLLIIVIASSL	GLNSGCAM			215					
GLPF_ECOLI	G	T	FSTYPNPHINFV	QAF	AVEMVITAILMGLILALTDDG	NGVPRGPLAPLLIGLLIAVIGASMG	PLTG	FAM		202					
ruler	...	180	...	190	...	200	...	210	...	220	...	230	...	240	...

The Aquaporin Superfamily

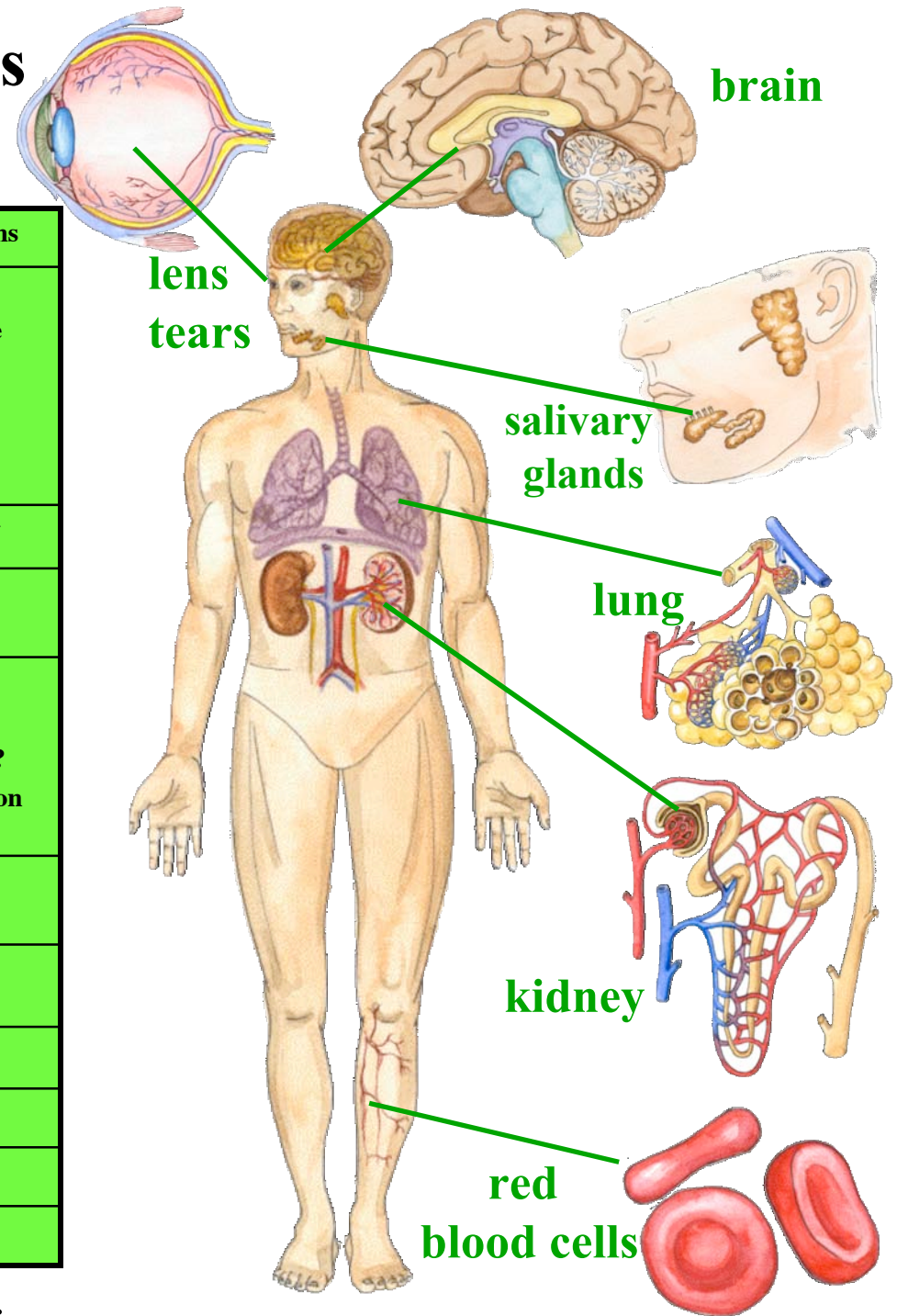
**Glycerol
transport**



Water and **Glycerol** Channels in the Human Body

Aquaporin-0	Eye: lens fiber cells	Fluid balance of the lens
Aquaporin-1	Red blood cells Kidney: proximal tubules Eye: ciliary epithelium Brain: choroid plexus Lung: alveolar epithelial cells	Osmotic protection Concentration of urine Aqueous humor Production of CSF Alveolar hydration
Aquaporin-2	Kidney: collecting ducts	ADH hormone activity
Aquaporin-3	Kidney: collecting ducts Trachea: epithelial cells	Reabsorption of water Secretion of water
Aquaporin-4	Kidney: collecting ducts Brain: ependymal cells Brain: hypothalamus Lung: bronchial epithelium	Reabsorption of water CSF fluid balance Osmosensing function? Bronchial fluid secretion
Aquaporin-5	Salivary glands Lacrimal glands	Production of saliva Production of tears
Aquaporin-6	Kidney	Very low water permeability!
Aquaporin-7	Testis and sperm	
Aquaporin-8	Testis, pancreas, liver	
Aquaporin-9	Leukocytes	
Aquaporin-10		

Additional members are suspected to exist.



Functionally Important Features of Aquaporins

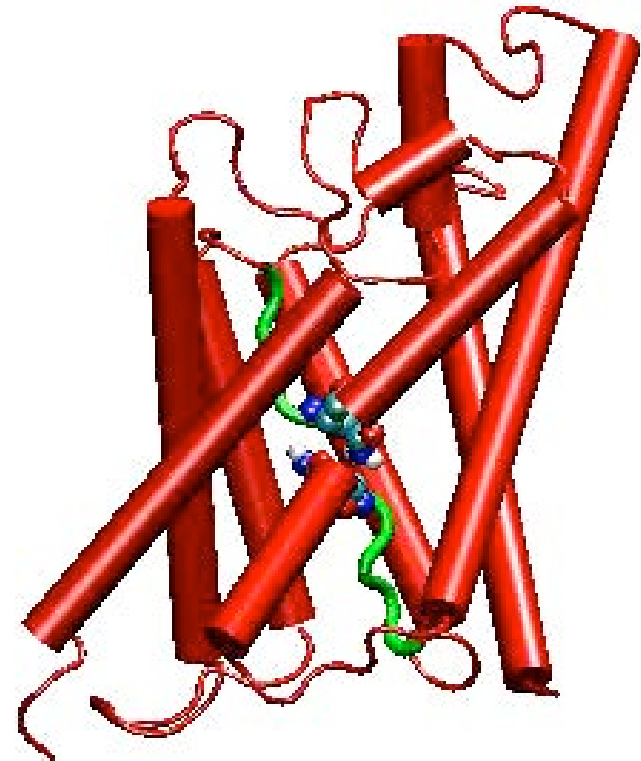
- Water and glycerol transport
- Exclusion of ions and protons
- Tetrameric arrangement in membrane

Aquaporins of known structure:

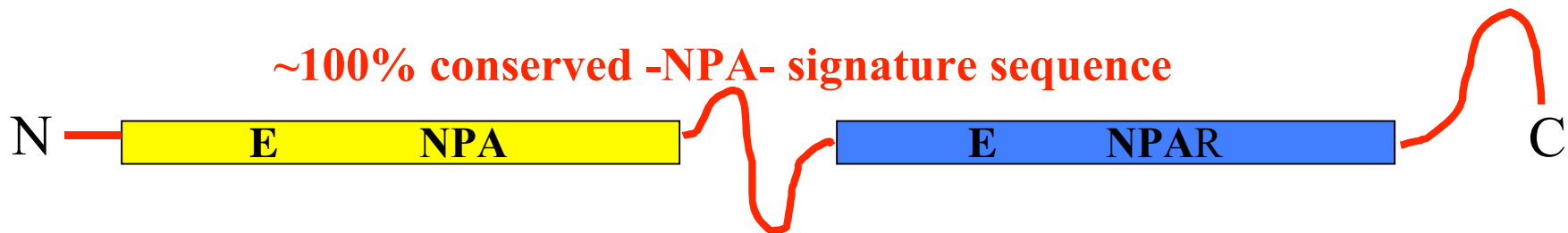
GlpF – E. coli glycerol channel (aquaglyceroporin)

– Fu, et al., Science (2000)

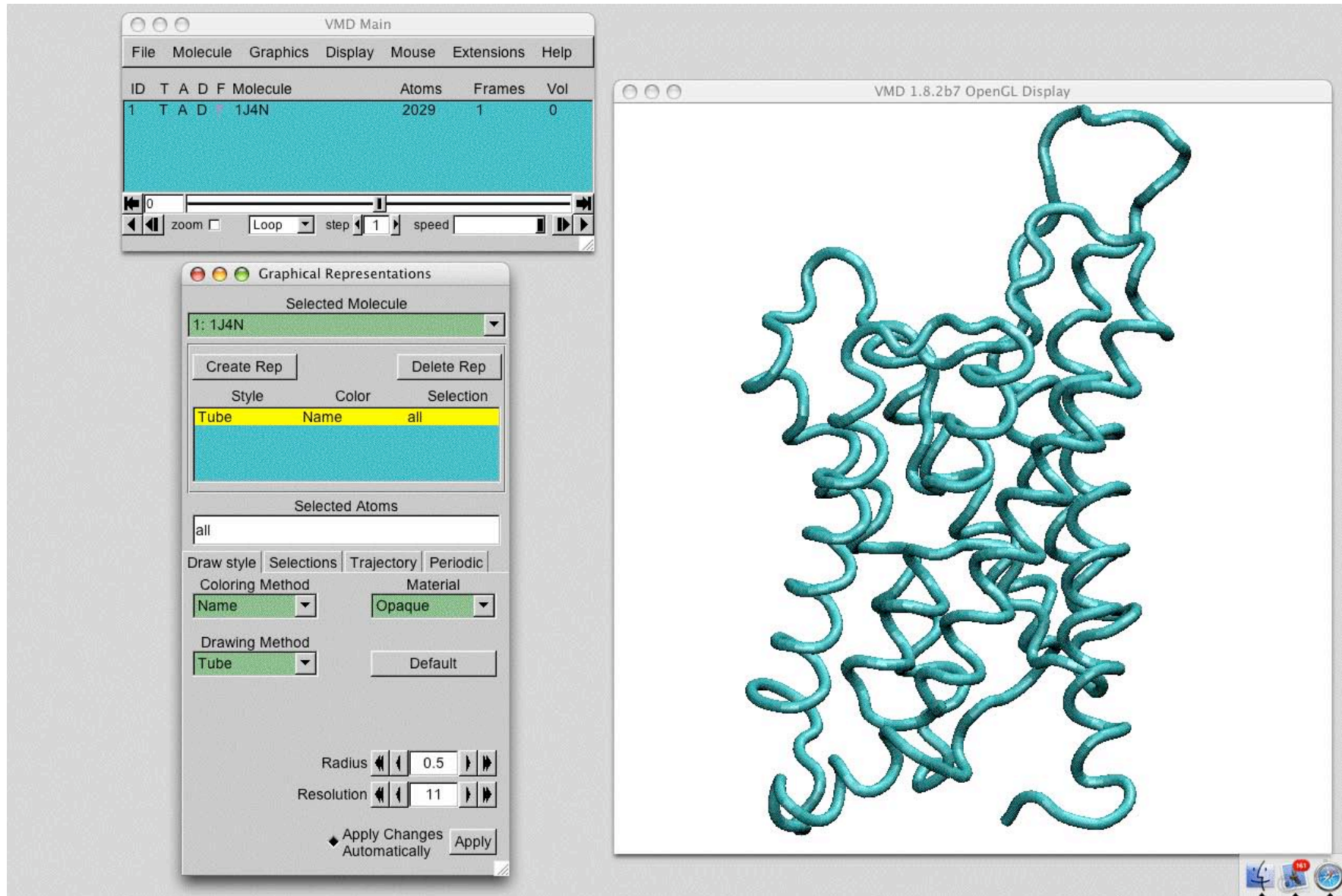
AQP1 – Mammalian aquaporin-1 (pure water channel) -Sui et al, Nature (2001)



~100% conserved -NPA- signature sequence

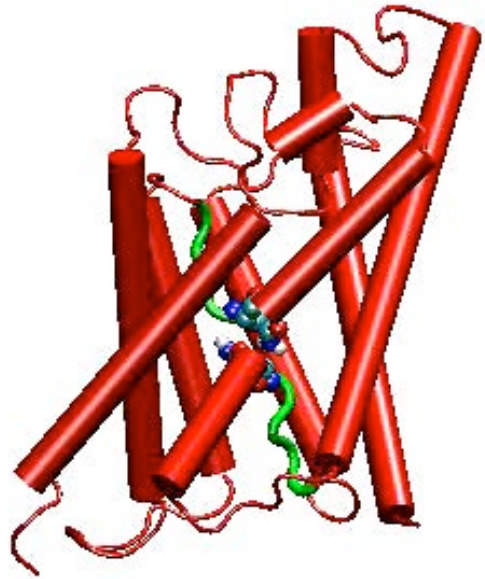


Load Aquaporin 1J4N into VMD

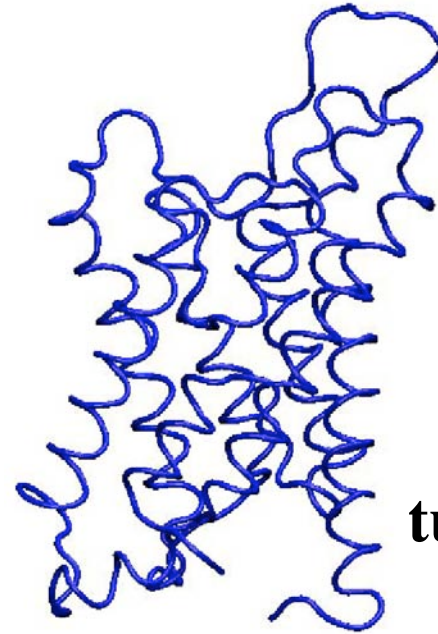


VMD Permits Different Rendering Styles

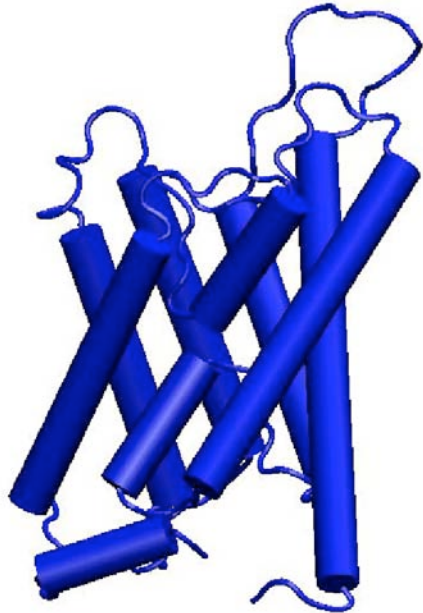
movie



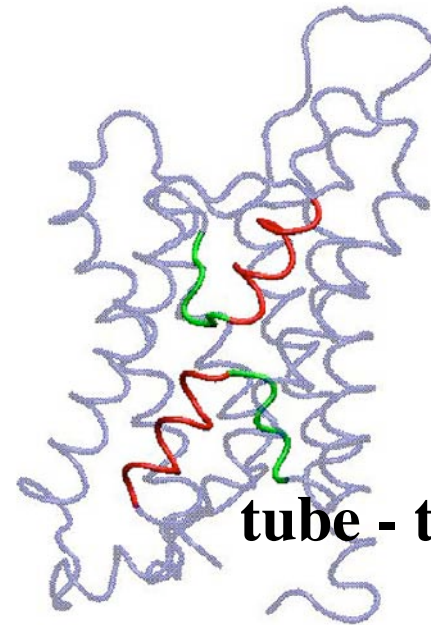
tube



cartoon



tube - transparent



Comparing Sequences of Aquaporins

```

      :      *  :  :  :  :  :
AQP0 HUMAN -----ASFWRAIFAFFAFLFYVFFGLGSSLRWA-----PGPLHVLQVAMAFGLALATLVSVGHI 62
AQP1 HUMAN -----KKLFWRRAVVAEFLATLLFVFISIGSALGFKYVPGNNQTAVQDNVKVSLAFGLSIATLAAVSGHI 70
AQP2 HUMAN -----IAFSRAVFAEFLATLLFVFGLGSAIWNF-----QALPSVLQIAMAFLGIGTLVQALGHI 62
AQP3 HUMAN -----IRYRLLRQALAECLGTLILVMFGCGSVAQVV-----LSRGTHGGFLTINLAFGFAVTLGILIAQV 77
AQP4 HUMAN -----VWTQAFWKAVTAEFAMLI FVLLSLGSTINWG-----GTEKPLPVDMLVLSLCFGLSIATMVQCFGHI 91
AQP5 HUMAN -----VAFLKAVFAEFLATLLFVFGLGSAIWNF-----SALPTILQIALAFGLAIGTLAALGPV 63
AQP6 HUMAN MLACRLWKATSRALFAEFLATGLYVFFGVGSVMRWPE-----TALPSVLQIAITFNILVTAMAVQVWKIT 76
AQP7 HUMAN KIQEILQRKMVREFLAEFMSTYVMMVFGLGSAVHNV-----LNK-KYGSYLGVNLFQFGVMTMGVHVAGRI 88
AQP8 HUMAN RWRVSWYERFVQPCLVELLGSAIFIFIGCLSVIENG-----TDTGLLPALAHGLALGLVIAITLGN 86
AQP9 HUMAN -----LKSSLAKETLSEFLGTFILIVLGCCEVAQAI-----LSRGRFGGVITINVGFSMAVAMAIYVAGGV 78
GLPF_ECOLI -----QTSTLKQCIAEFLGTGLLIFFGVGCVAALK-----VAGASFG-QWEISVIWGLGVAMAIYLTAGV 62
ruler      ...40.....50.....60.....70.....80.....90.....100....

```



```

      **.* *.* :  :  :  :  :
AQP0 HUMAN SGAHVNPAVTFAFLVGSQMSLLRAFCYMAAQLLGAVAGAALVSVTP-----PAVRGNLA----- 117
AQP1 HUMAN SGAHLNPAVTLGLLLSQCIISIFRALMYIIAQCQVGAIVATAILSGITS-----SLTGNSLG----- 125
AQP2 HUMAN SGAHINPAVTVACLVGCHVSVLRAAFYVAAQLLGAVAGAALLHEITP-----ADIRGDLA----- 117
AQP3 HUMAN SGAHLNPAVTTFAMCFLAREPWIKLPIYTLAQTLAGFLGAGIVFGLYDAIWHFADNQLFVSGPNG-----TA 144
AQP4 HUMAN SGGHINPAVTVMVCTRKISIAKSVFYIAAQCGLGAIIGAGILYLVTF-----PSVVGGLG----- 146
AQP5 HUMAN SGGHINPAITLALLVGNQISLLRAFFVYVAAQLVGAIAAGAGILYGVAP-----LNARGNLA----- 118
AQP6 HUMAN SGAHANPAVTLAFLVGSHTSLPRAVAVVAAQLVGATVGAALYGVMP-----GDIRETLG----- 131
AQP7 HUMAN SGAHMNAAVTFANCALGRVPWRKFPPVVLGQFLGSFLAAATISLFTYAILHFSGGQLMVTGPVA-----TA 155
AQP8 HUMAN SGGHFNPAVSLAAMLIIGGLNLVMLLPYVWSQLLGGMLGAALAKVVS-----ERFVWNASG----- 142
AQP9 HUMAN SGGHINPAVSLAMCLFGRMKWFKLPYVVGAFVGAQFLGAFVGAATVFGIYDGLMSFAGGKLLIVGENA-----TA 145
GLPF_ECOLI SGAHLNPAVTIALWLWFACFDKRKVLPFIVSQVAGAFCAAALVYGLYVNLFFDFEQTHIVRGSVEVDLA 132
ruler      ...110.....120.....130.....140.....150.....160.....170....

```



```

      :      *  :  :  :  :
AQP0 HUMAN ---LNTLHPAVSVGQATTVEIFLTLQVLCIFATYDE--RRNQQLGSSVALAVGFSALGHLFGMYTGLAGM 183
AQP1 HUMAN ---RNDLADGVNSGGGLGIBIIGTLQVLCVLAT--TDR--RRDLGGSAPLAIGLSVALGHLLAIDYTGCGI 191
AQP2 HUMAN ---VNALSNSITAGQAVTVLFLTLQVLCIFASTDE--RRGENPGTPALSIGFSVALGHLLGIHYTGCSM 183
AQP3 HUMAN GIFATYPSGHLDMINGFFDQFIGTASLIVCVLAIVDPYNNNVPVRGLEAFVGLVVLVIGTSMGFNSGYAV 214
AQP4 HUMAN ---VTMVHGNLTAGHGLLVELIITFLVFTIFASCD--KRTDVTGSIALAIGFSVAIGHLFAINYTGASM 212
AQP5 HUMAN ---VNALNNNTTQGOAMVVELILTLQALCIFASTDS--RRTSPVGSPPALSIGLSVTLGHLVGIYFTGCSM 184
AQP6 HUMAN ---INVVRNSVSTGQAVAVELLTLQVLCVFASTDS--RQTS--GSPATMIGISWALGHLIGILFTGCSM 195
AQP7 HUMAN GIFATYLPDHMTLWRGFLNEAWLTGMLQLCLFATIDQENNPAIPGTEALVIGILVVTIGVSLGMNTGYAI 225
AQP8 HUMAN -AAFVTVQEQGVAGALVABEILTLALAVCMGAIN--EKTGKPLAPFSIGFAVTVDILAGGPVSGGCM 209
AQP9 HUMAN HIFATYPPAPYLSLANAFADQVATMILLIIVFAIFDSRNLGAPRGLEPTAIGLLIIVIASSLGLNSGCAM 215
GLPF_ECOLI GTFSTVPNPHINFVQAFVEMVITAILMGLILALDDGNGVPRGPLAPLLIGLLIAVIGASMGPLTGFAM 202
ruler      ...180.....190.....200.....210.....220.....230.....240....

```



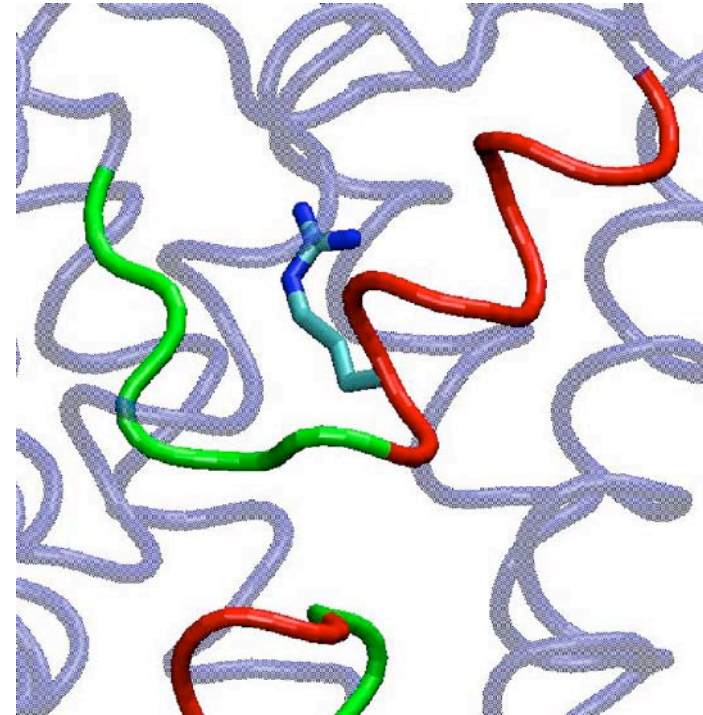
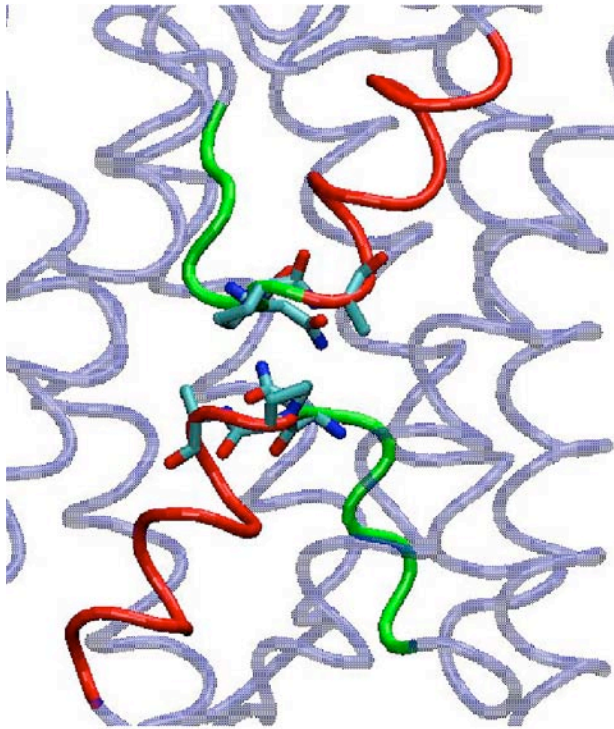
```

      *** :  :  :  :
AQP0 HUMAN NPARSFAPAILTGNFT-----NHWVYVVGPIIGGGLGSLLYDFLLFP----- 225
AQP1 HUMAN NPARSFGSAVITHNFS-----NHWIFWVGPFIGGALAVLIYDFILAP----- 233
AQP2 HUMAN NPARSLAPAVVTGKFD-----DHWVFWIGPLVGAILGSLLYVNLVLP----- 225
AQP3 HUMAN NPARDFGPRFLTALAGWSAVFTTGG--HWWVVPVPSLLGSTIAGVFVYQLMIGC----- 267
AQP4 HUMAN NPARSFGPAVIMGNWE-----NHWIYVVGPIIGAVLAGGLYEVVFCPD----- 255
AQP5 HUMAN NPARSFGPAVVMNRFSP-----AHWVFWVGPIVGAVALAAILYFLLFP----- 227
AQP6 HUMAN NPARSFGPAIITGKFT-----VHWVFWVGPLMGALLASLIYNFVLP----- 237
AQP7 HUMAN NPSRDLFPRIFTFIAGWGKQVFSNGE--NWWWVPVVPVAPLLGAYLGGIIVLVFIS----- 278
AQP8 HUMAN NPARAFGPAVVANHNW-----FHWIYVWLGPLLGLLVGLLIRCFID----- 251
AQP9 HUMAN NPARDLSPRLFTALAGWGFVFRAGN--NFWWIYVVGPLVGAVIGGLIYVLVIEI----- 268
GLPF_ECOLI NPARDFGPKVFAWLAWGNVAFITGGRDIPYFLVPLFGPIVGAIVGAFARKLIGR----- 257
ruler      ...250.....260.....270.....280.....290.....300....

```



Highlighting Key Conserved Residues



		*** :	..			:	..*::..		:																																												
AQP0	HUMAN	NP	ARS	F	PA	IL	T	GN	FT	-----NHWVYVWGPII	GGGL	GS	LL	YD	FL	LF	225																																				
AQP1	HUMAN	NP	ARS	F	GS	AV	I	TH	NS	-----NHWIFWVGPII	GGAL	AV	LI	YD	FI	LP	233																																				
AQP2	HUMAN	NP	ARSL	AP	AV	VT	GK	FD	-----DHVWFVIGPLV	GAIL	GS	LL	YN	YV	LF	225																																					
AQP3	HUMAN	NP	ARD	F	G	P	R	L	F	TALAGWGS	AV	F	T	T	GQ	-----HWWVWPIV	SPLL	GS	IA	GV	FV	YQ	LM	IG	267																												
AQP4	HUMAN	NP	ARS	F	G	P	AV	IM	GN	WE	-----NHWIYVWGPII	GAVL	AG	GL	YE	V	FC	PD	255																																		
AQP5	HUMAN	NP	ARS	F	G	P	AV	VM	NR	FS	-----AHVWFVWGPIV	GAVL	AA	IL	YF	YL	LF	227																																			
AQP6	HUMAN	NP	ARS	F	G	P	AI	I	G	K	FT	-----VHVWFVWGPLM	GALL	AS	LI	YN	FV	LF	237																																		
AQP7	HUMAN	NP	SR	D	L	P	P	R	I	F	T	FI	AG	W	G	K	Q	V	F	S	N	G	E	-----NWWVWPV	VAPL	L	G	AY	L	G	GI	I	Y	L	V	F	I	G	S	278													
AQP8	HUMAN	NP	AR	A	F	G	P	AV	V	A	N	H	W	N	-----FHWIYWL	G	P	L	L	A	G	L	L	V	G	L	L	I	R	C	F	I	G	D	251																		
AQP9	HUMAN	NP	ARD	L	S	P	R	L	F	TALAGW	G	F	E	V	F	R	A	G	N	-----NFWWI	P	V	V	G	P	L	V	G	A	V	I	G	G	L	I	Y	V	L	V	I	E	I	268										
GLPF	ECOLI	NP	ARD	F	G	P	K	V	F	A	W	L	A	G	W	G	N	V	A	F	T	G	G	R	D	I	P	Y	F	L	V	P	L	F	G	P	I	V	G	A	I	V	A	F	A	Y	R	K	L	I	G	R	257
	ruler	250	260	270	280	290	300																															



Load Aquaporins 1j4n, 1fqy, 1lda, 1rc2 into VMD

The screenshot displays the VMD (Visual Molecular Dynamics) software interface. The main window, titled "VMD Main", shows a list of loaded molecules:

ID	T	A	D	F	Molecule	Atoms	Frames	Vol
1		A	D		1J4N	2029	1	0
2		A	D		1FQY	1661	1	0
3		A	D		1lda	1997	1	0
5	T	A	D		1rc2	3530	1	0

The "Graphical Representations" window shows the selected molecule "5: 1rc2" with a "Tube" style, "ColorID 3" color, and "chain A" selection. The "Drawing Method" is set to "Tube".

The "VMD 1.8.2b7 OpenGL Display" window shows a 3D ribbon representation of the protein structure, colored by chain (red, blue, yellow, and orange).

The "Multiple Sequence Alignment" window displays the following sequence alignment:

PDB code	Description
1j4n	Bovine AQP1
1fqy	Human AQP1
1lda	E. coli Glycerol Facilitator (GlpF)
1r2c	E. coli AqpZ

The sequence alignment shows the following sequences:

```
d1fqa_.ent KLPWRAVVAEFLATTLFVFISIGSALGFKYPVGNWQTAVQDNVKSIAFGLSIATLA
d1j4na_.ent HASEFKKKLFWRAVVAEFLAHILFIPISIGSALGFHYPIKSNQTTGAVQDNVKSIA
d1lda_.ent TLNGQCIAEFLDTGLLIFFGVGCVAAKVAGASFGQWEISVINGLGVMATYLTAGV
d1rc2a_.ent HFRKLAEECFGTFWLVFGGCGSAVLAAGFPFELGIGFAGVALAFGLTVLTHAFVGH
```


Aligning Structures and Sequences

The image displays the VMD (Visual Molecular Dynamics) software interface, which is used for visualizing and analyzing molecular structures and sequences. The interface is divided into several windows:

- VMD Main:** This window contains a menu bar (File, Molecule, Graphics, Display, Mouse, Extensions, Help) and a table of loaded molecules. The table lists four molecules: 1J4N, 1FQY, 1lda, and 1rc2. Below the table is a playback control bar with buttons for zoom, loop, step, and speed.
- Graphical Representations:** This window allows users to manage the visual representation of the selected molecule (5: 1rc2). It includes buttons for 'Create Rep' and 'Delete Rep', a table for selecting atoms (chain A), and options for coloring (ColorID 3) and drawing (Tube) methods. It also features a 'Default' button and a checkbox for 'Apply Changes Automatically'.
- VMD 1.8.2b7 OpenGL Display:** This window shows a 3D visualization of the protein structure, rendered as a ribbon model. The structure is colored by chain (chain A is yellow) and is shown in a transparent tube representation.
- Multiple Sequence Alignment:** This window displays a multiple sequence alignment of the four protein sequences. The sequences are aligned based on their primary structure, and the alignment is shown in a text-based format. The sequences are: 1fgy, 1j4n, 1lda, and 1rc2. The alignment shows conserved regions across the sequences.

ID	T	A	D	F	Molecule	Atoms	Frames	Vol
1		A	D	F	1J4N	2029	1	0
2		A	D	F	1FQY	1661	1	0
3		A	D	F	1lda	1997	1	0
5	T	A	D	F	1rc2	3530	1	0

Selected Molecule: 5: 1rc2

Create Rep Delete Rep

Style	Color	Selection
Tube	ColorID 3	chain A

Selected Atoms: chain A

Draw style: Selections Trajectory Periodic

Coloring Method: ColorID 3 Material: Opaque

Drawing Method: Tube Default

Radius: 0.5 Resolution: 11

Apply Changes Automatically Apply

Multiple Sequence Alignment

Align Molecules... FASTA Highlight PDB Pairwise RMSD Sequence Display

```
1fgy -----KLFWRVVAEFLATTLFVFTSIGSAL-GF-KY---PVGNNQTAVQDNVKVSLAPGLSIATLAQS-VGHISGAHLNPAVTLGILLSCQISIF-RAI
1j4n MASEFKKKLFWRVVAEFLAMILFIFTISIGSAL-GF-HYPIKENQT-TGAVQDNVKVSLAPGLSIATLAQSVGH-ISGAHLNPAVTLGILLSCO-ISVLRAI
1lda -----TLKGQCIAEFLGTGLLIFFGVGCVA-ALKVA-----G-A-SFGQWEISVINGLGVAMAIYLTAGVSGAHLNPAVTIALWLFA-CFDKRVV
1rc2 -----MFRKLAAECFGTFWLVFGCCGSAVLA-AG-----FPE-LGIGFAGVALAPGLTVLTMFAVVG-HISGGHFNPAVTIGLNAGG-RFPKREV
```

Comparing Structures by Similarity - Q Value

The image displays the VMD (Visual Molecular Dynamics) software interface, which is used for visualizing molecular structures and performing sequence alignments.

VMD Main Window: This window shows a list of loaded molecules. The table below represents the data shown in the interface:

ID	T	A	D	F	Molecule	Atoms	Frames	Vol
1		A	D		1J4N	2029	1	0
2		A	D		1FQY	1661	1	0
3		A	D		1lda	1997	1	0
5	T	A	D		1rc2	3530	1	0

Graphical Representations Window: This window allows users to configure the visualization of the selected molecule (5: 1rc2). The "Style" is set to "Tube", "Color" is "ColorID 3", and "Selection" is "chain A". The "Drawing Method" is also set to "Tube".

Multiple Sequence Alignment Window: This window displays a sequence alignment of four proteins: 1f4y, 1j4n, 1lda, and 1rc2. A context menu is open, showing options for "Molecule Coloring" and "Highlight Style". The "Q per residue" option is selected under "Molecule Coloring".

OpenGL Display Window: This window shows a 3D visualization of the protein structure, rendered as a multi-colored ribbon model, representing the spatial arrangement of the atoms.

Comparing Structures by Similarity - Q Value

The image displays the VMD (Visual Molecular Dynamics) software interface, which is used for visualizing molecular structures and performing sequence alignments.

VMD Main Window: This window shows a list of loaded molecules. The table below represents the data shown in the interface:

ID	T	A	D	F	Molecule	Atoms	Frames	Vol
1		A	D		1J4N	2029	1	0
2		A	D		1FQY	1661	1	0
3		A	D		1lda	1997	1	0
5	T	A	D		1rc2	3530	1	0

Graphical Representations Window: This window allows users to configure the visualization of the selected molecule (5: 1rc2). The "Style" is set to "Tube", the "Color" is "ColorID 3", and the "Selection" is "chain A". The "Drawing Method" is also set to "Tube".

VMD 1.8.2b7 OpenGL Display Window: This window displays the 3D visualization of the protein structure, showing multiple chains in different colors (red, green, blue, yellow) to represent different residues or regions.

Multiple Sequence Alignment Window: This window shows a sequence alignment of four proteins: 1f4y, 1j4n, 1lda, and 1rc2. The alignment is displayed as a grid of characters. A context menu is open, showing options for "Molecule Coloring" and "Highlight Style". The "Q per residue" option is selected under "Molecule Coloring".

Exhibiting Sequence Identity - Side View

The image displays the VMD (Visual Molecular Dynamics) software interface, showing a protein structure in a side view and a sequence alignment window.

VMD Main Window:

ID	T	A	D	F	Molecule	Atoms	Frames	Vol
1		A	D	F	1J4N	2029	1	0
2		A	D	F	1FQY	1661	1	0
3		A	D	F	1lda	1997	1	0
5	T	A	D	F	1rc2	3530	1	0

Graphical Representations Window:

Selected Molecule: 5: 1rc2

Create Rep Delete Rep

Style	Color	Selection
Tube	ColorID 3	chain A

Selected Atoms: chain A

Draw style: Selections Trajectory Periodic

Coloring Method: ColorID 3 Material: Opaque

Drawing Method: Tube Default

Radius: 0.5 Resolution: 11

Apply Changes Automatically Apply

VMD 1.8.2b7 OpenGL Display Window:

Shows a 3D ribbon representation of the protein structure, colored by sequence identity (red, green, blue, yellow).

Multiple Sequence Alignment Window:

Align Molecules... FASTA Highlight PDB Pairwise RMSD Sequence Display

Sequence alignment showing residues 1f4y, 1j4n, 1lda, and 1rc2. Residues are highlighted in yellow to indicate sequence identity.

Exhibiting Sequence Identity - Top View

The screenshot displays the VMD (Visual Molecular Dynamics) software interface, illustrating the visualization of protein structure and sequence identity.

VMD Main Window: Shows a list of loaded molecules:

ID	T	A	D	F	Molecule	Atoms	Frames	Vol
1		A	D	F	1J4N	2029	1	0
2		A	D	F	1FQY	1661	1	0
3		A	D	F	1lda	1997	1	0
5	T	A	D	F	1rc2	3530	1	0

Graphical Representations Window: Configures the display of the selected molecule (5: 1rc2).

- Selected Molecule:** 5: 1rc2
- Create Rep / Delete Rep:** Buttons to manage representations.
- Style / Color / Selection:** Table showing representation details.
- Selected Atoms:** chain A
- Draw style:** Selections, Trajectory, Periodic
- Coloring Method:** ColorID 3
- Material:** Opaque
- Drawing Method:** Tube
- Radius:** 0.5
- Resolution:** 11
- Apply Changes Automatically:** Checkmark
- Apply:** Button

VMD 1.8.2b7 OpenGL Display Window: Shows the 3D visualization of the protein structure (1rc2) as a ribbon model, colored by sequence identity (ColorID 3).

Multiple Sequence Alignment Window: Displays the sequence alignment of the four proteins (1f4y, 1j4n, 1lda, 1rc2) with highlighted regions indicating sequence identity.

```
1f4y -----XLFWRVAVAEFLATTLFVFTISIGSAL-GF-KY---FVGNQTAVDNWKVSLAFGLSIATLAQS-VGHIISGAHLNPAVTLGLLLSCQISIF-RV
1j4n MASEPKKKLFWRAVVAEFLAMILFIFISIGSAL-GF-HYPIKSNQ-TGAVQDNVKSLSAFGLSIATLAQSVGH-ISGAHLNPAVTLGLLLSCO-ISVLRV
1lda -----TLKGQCIAEFLGTGLLFTFGVGVVA-ALKVA-----G-A-SFGQWEISVINGLVAMAIYLTA-GVSGAHLNPAVTIALWLFA-CFDKRV
1rc2 -----MFRKLAEECFOTFWLVPFGCCSAULA-AG-----FPE-LGIGFAGVALAFGLTVLTMAFAVG-HISGGHNPVAVTIGLWAGG-RFPAREV
```

Showing Conserved Residues - Monomer

The screenshot displays the VMD (Visual Molecular Dynamics) software interface, which is used for visualizing molecular structures and performing sequence alignments.

VMD Main Window: This window shows a list of loaded molecules. The table below summarizes the data:

ID	T	A	D	F	Molecule	Atoms	Frames	Vol
1		A	D		1J4N	2029	1	0
2		A	D		1FQY	1661	1	0
3		A	D		1lda	1997	1	0
5	T	A	D		1rc2	3530	1	0

Below the table are playback controls including a timeline, zoom, loop, step, and speed settings.

Graphical Representations Window: This window allows for customizing the display of the selected molecule (5: 1rc2). The 'Selected Molecule' dropdown is set to '5: 1rc2'. The 'Create Rep' button is visible. The 'Style' dropdown is set to 'Tube', 'Color' is 'ColorID 3', and 'Selection' is 'chain A'. The 'Selected Atoms' field contains 'chain A'. Under the 'Draw style' tab, 'Coloring Method' is 'ColorID', 'Material' is 'Opaque', and 'Drawing Method' is 'Tube'. At the bottom, 'Radius' is set to 0.5 and 'Resolution' is set to 11. The 'Apply Changes Automatically' checkbox is checked.

VMD 1.8.2b7 OpenGL Display Window: This window shows a 3D ribbon representation of the protein structure. The protein is colored by chain, with chain A highlighted in yellow and other chains in red and green.

Multiple Sequence Alignment Window: This window displays a sequence alignment of four proteins: 1fgy, 1j4n, 1lda, and 1rc2. The alignment is shown in a text-based format with yellow highlighting indicating conserved residues. The sequence display is as follows:

```
1fgy  -----XLFWRVAVAEFLATTLFVFISIGSAL-GF-KY---PVGNQTAVDNWKVSLAFGLSIATLAQS-VGHSAGHLNPAVTLGLLLSCQISIF-RV
1j4n  MASEPKKKLFWRAVVAEFLAMILFIFISIGSAL-GF-HYPIKSNQ-TGAVQDNVKVSLAFGLSIATLAQSVGH-ISGHLNPAVTLGLLLSCO-ISVLRV
1lda  -----TLRGQCIAEFLGTGLLFTGVGCVV-ALKVA-----G-A-SFGQWEISVINGLVAMAIYLTA-GVSGHLNPAVTIALWLFA-CFDKRV
1rc2  -----MFRKLAEECPQTFWLVFGCCSAVLA-AG-----FPE-LGIGPAGVALAFGLTVLTMAFAVG-HISGGHNPVAVTIGLWAGG-RFPAREV
```


Showing Conserved Residues - Tetramer

VMD Main

ID	T	A	D	F	Molecule	Atoms	Frames	Vol
1		A	D		1J4N	2029	1	0
2		A	D		1FQY	1661	1	0
3		A	D		1lda	1997	1	0
5	T	A	D		1rc2	3530	1	0

0 [progress bar] [play] [stop] [reset]

zoom ☐ Loop step 1 speed [slider]

Graphical Representations

Selected Molecule
5: 1rc2

Create Rep Delete Rep

Style	Color	Selection
Tube	ColorID 3	chain A

Selected Atoms
chain A

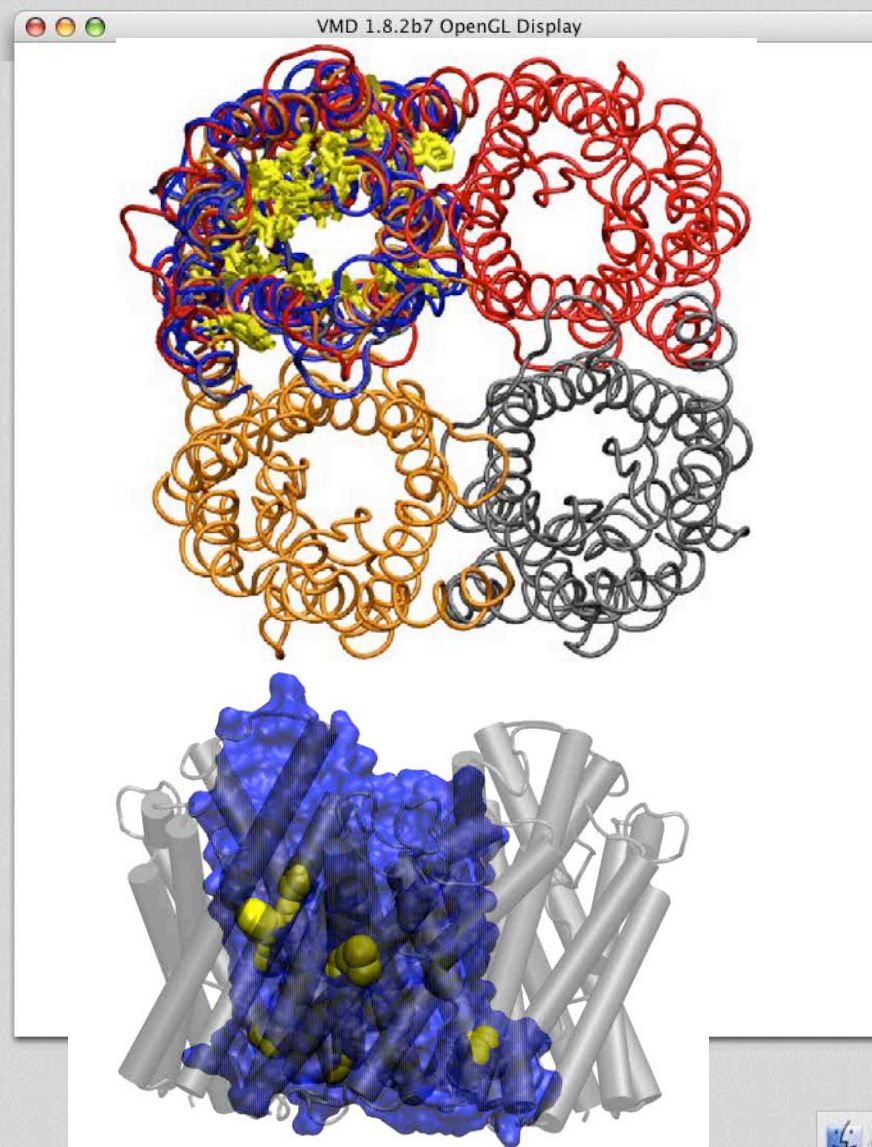
Draw style Selections Trajectory Periodic

Coloring Method
ColorID 3 Material Opaque

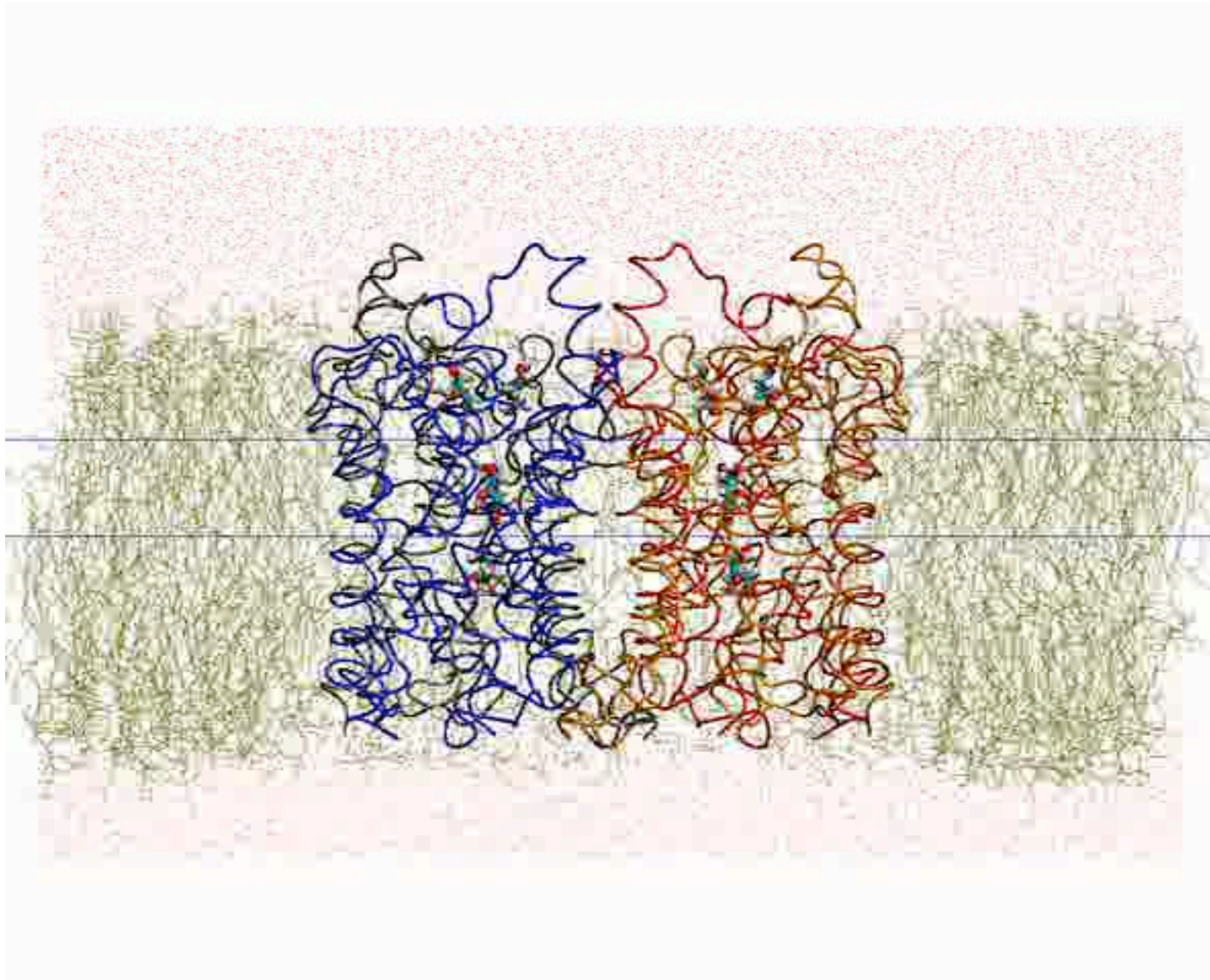
Drawing Method
Tube Default

Radius 0.5 Resolution 11

Apply Changes Automatically Apply

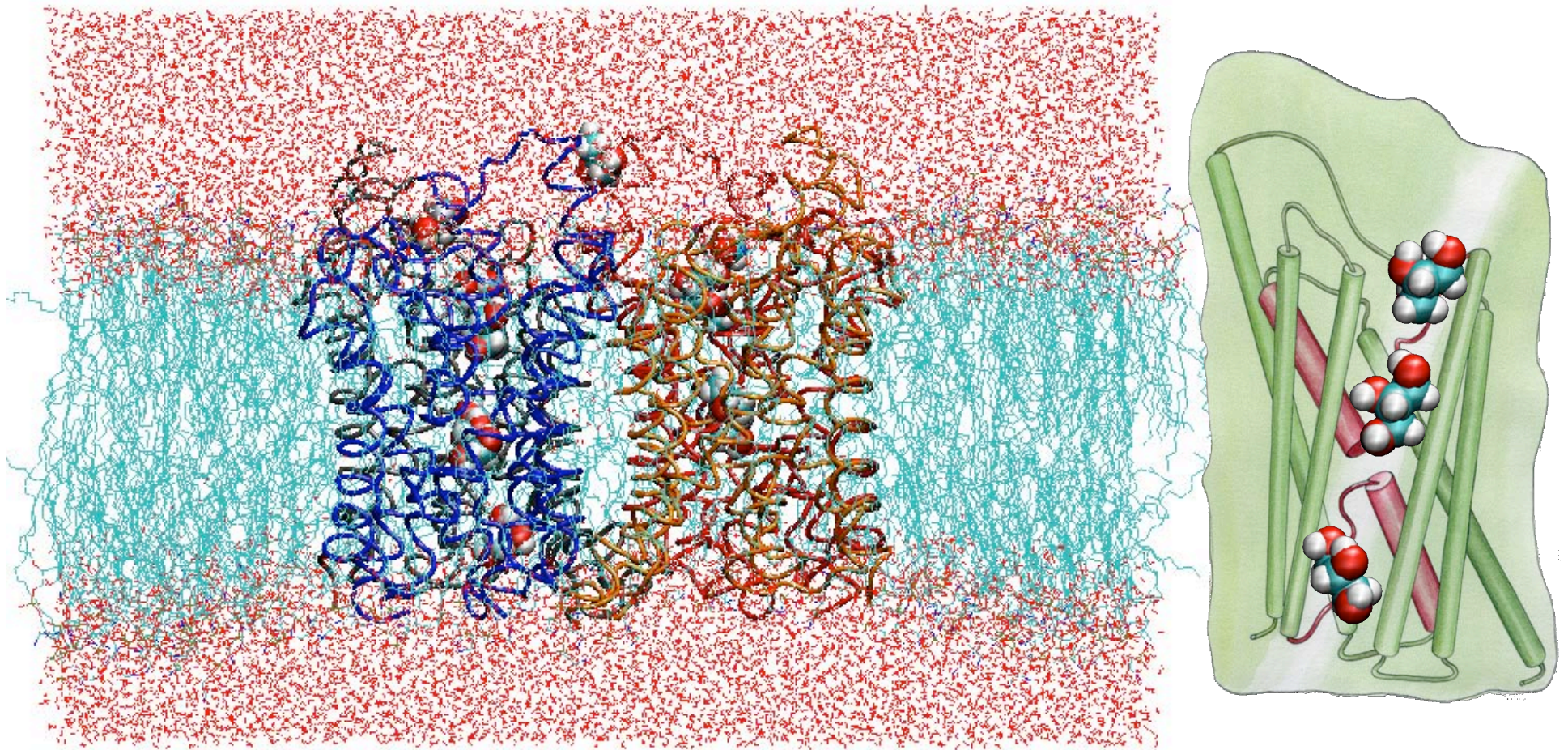


Dynamics of Protein, Lipid, Water System



M. Jensen, E. Tajkhorshid, K. Schulten, *Structure* 9, 1083 (2001)

Equilibrated Structure after 1 ns



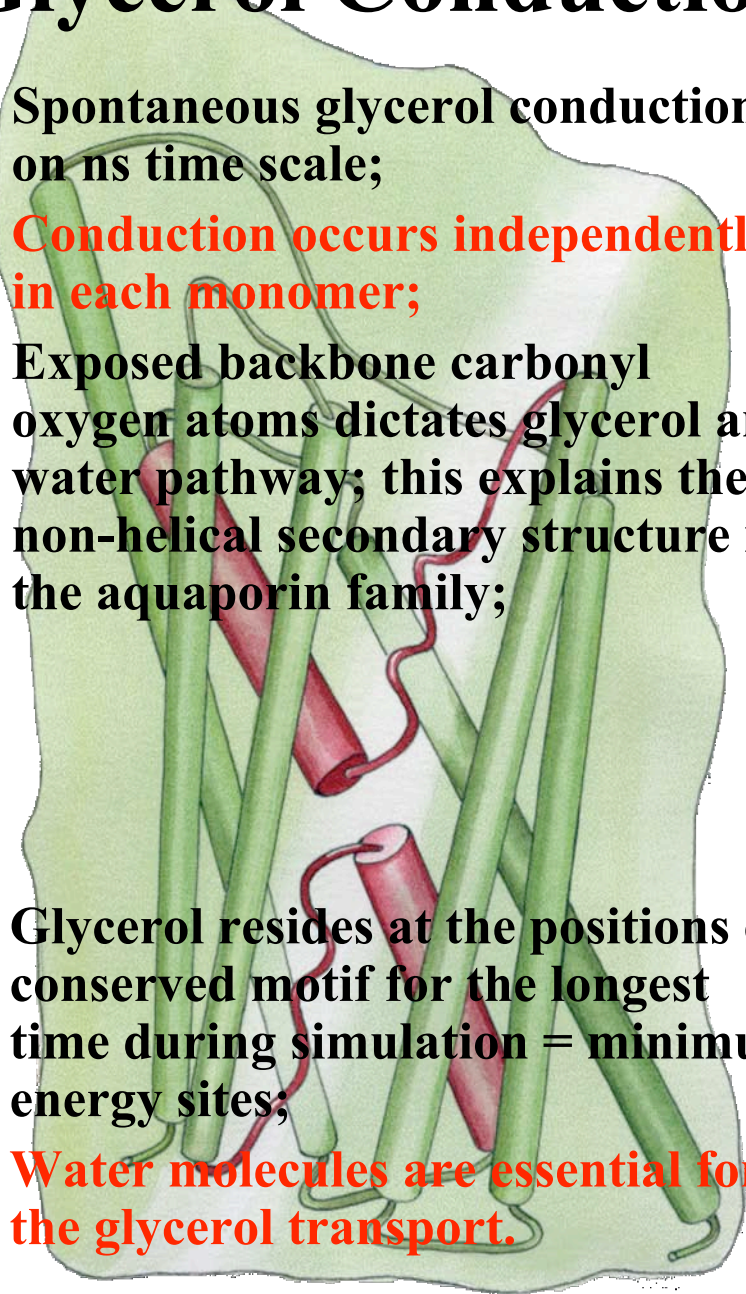
note the curved adjustment between lipids-protein

Morten Jensen, Emad Tajkhorshid

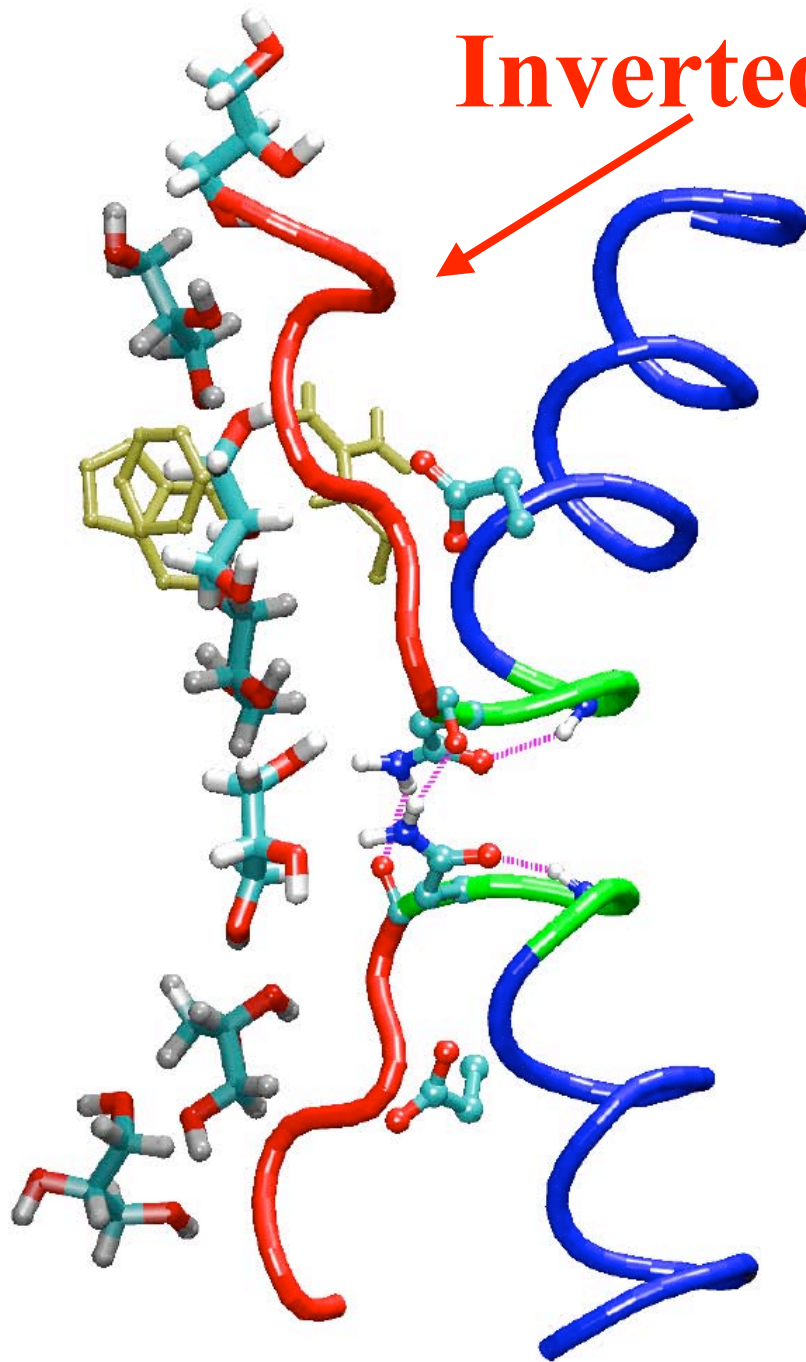
Glycerol Conduction



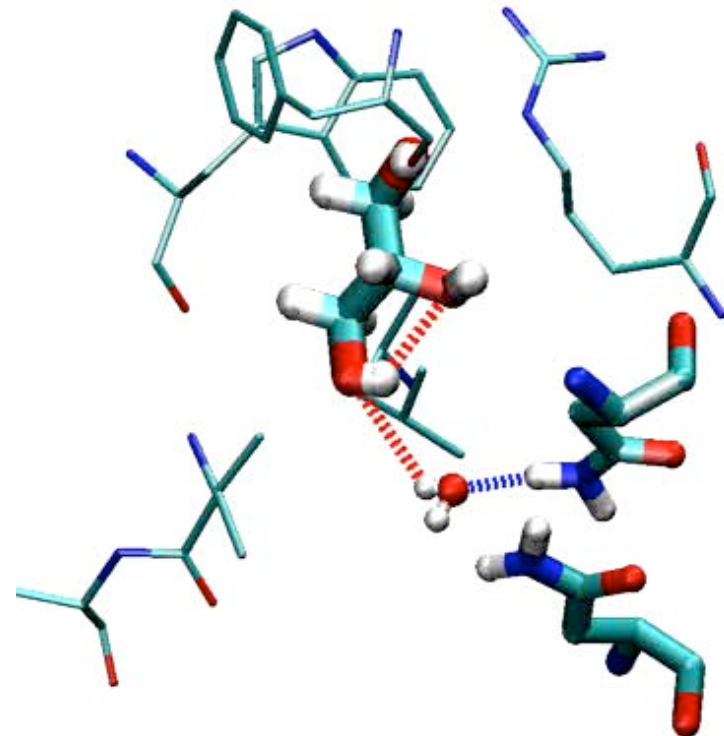
- Spontaneous glycerol conduction on ns time scale;
- **Conduction occurs independently in each monomer;**
- Exposed backbone carbonyl oxygen atoms dictates glycerol and water pathway; this explains the non-helical secondary structure in the aquaporin family;
- Glycerol resides at the positions of conserved motif for the longest time during simulation = minimum energy sites;
- **Water molecules are essential for the glycerol transport.**



Inverted helices guide glycerol



Glycerol – water competition for hydrogen bonds drives transport



Interactive Molecular Dynamics

VMD ←.....→ **NAMD**

Molecular Graphics

Molecular Dynamics

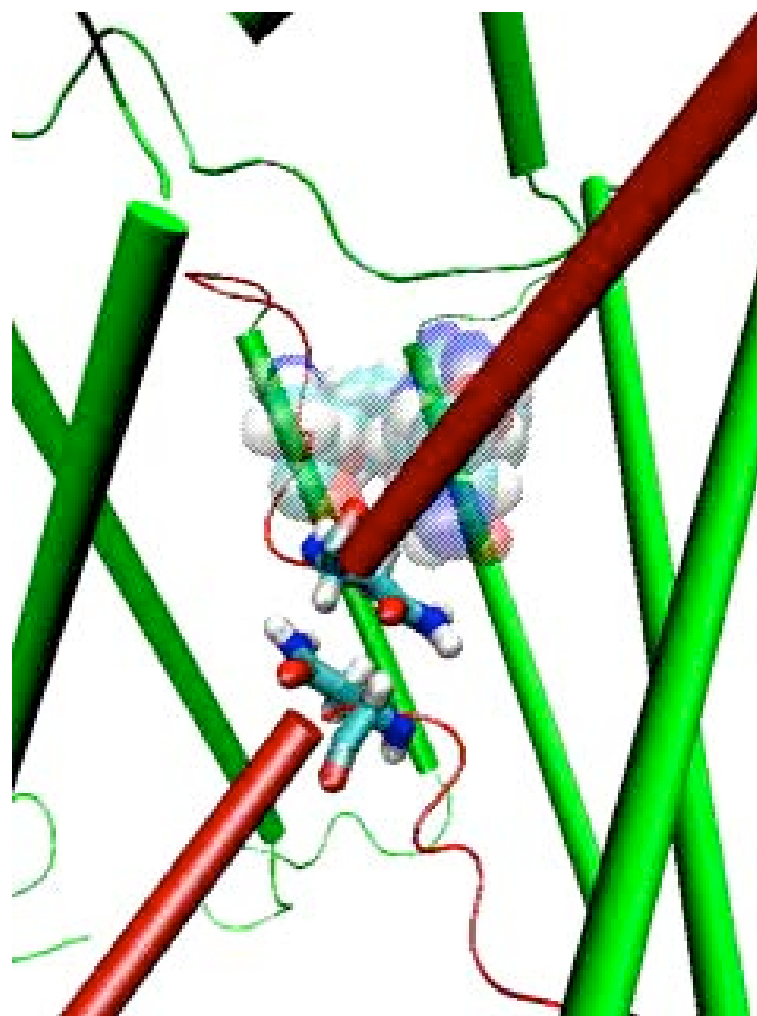
- Any PC/Workstation
- Supports 3D force-feedback devices for interaction



J. Stone, J. Gullingsrud, K. Schulten, and P. Grayson.
A System for Interactive Molecular Dynamics Simulation.
2001 ACM Symposium on Interactive 3D Graphics,
pp.191-194, ACM SIGGRAPH

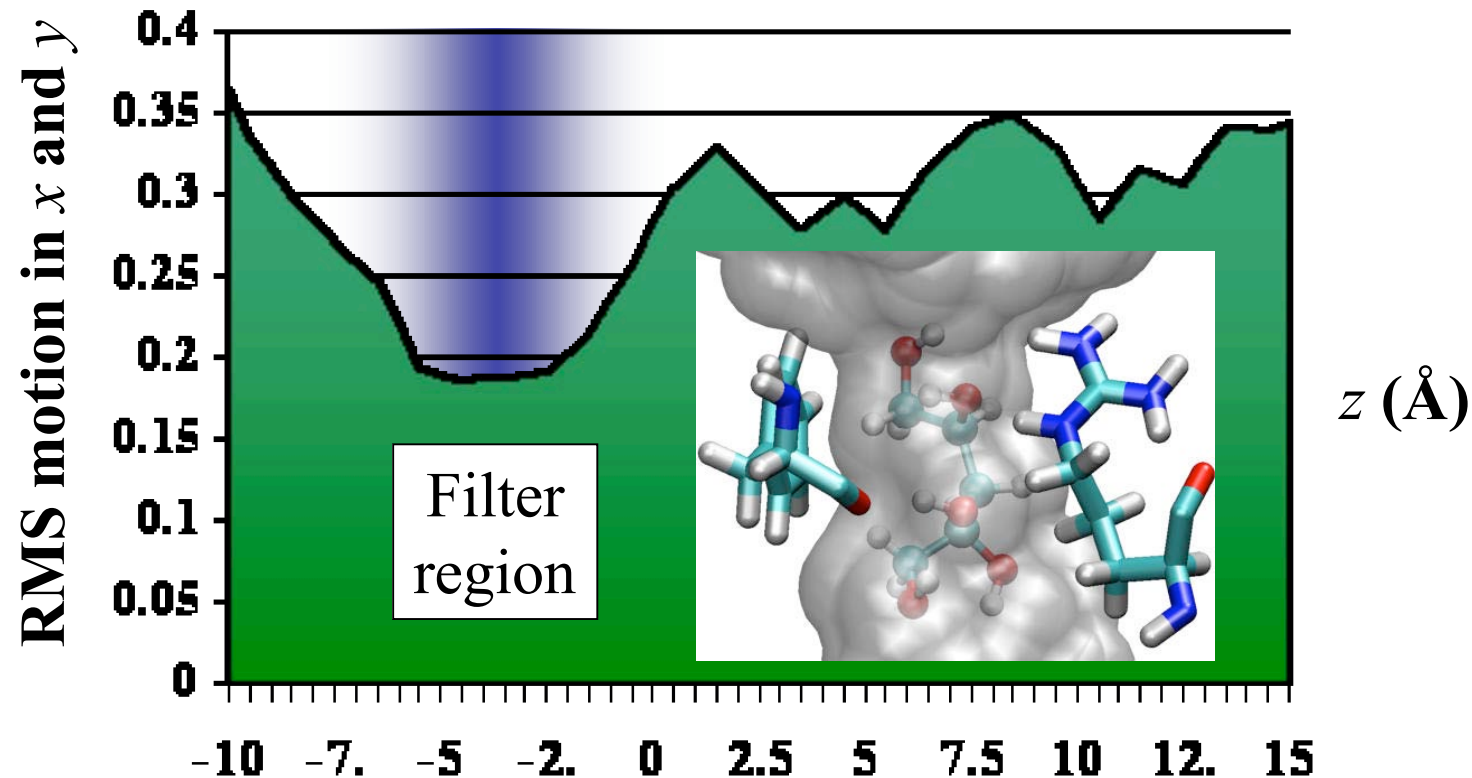
P. Grayson, E. Tajkhorshid, and K. Schulten.
Biophysical J, **83**: 36-48 (2003)

NIH Resource for Macromolecular Modeling and Bioinformatics
Theoretical Biophysics Group, Beckman Institute, UIUC



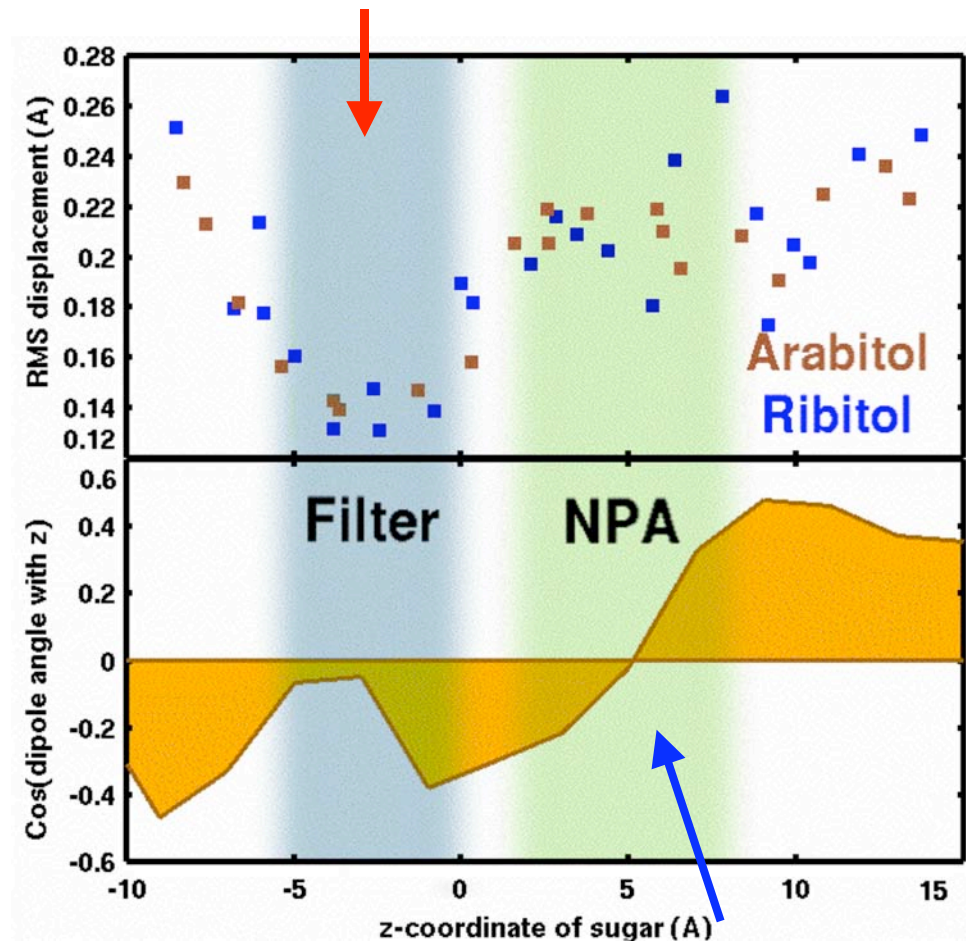
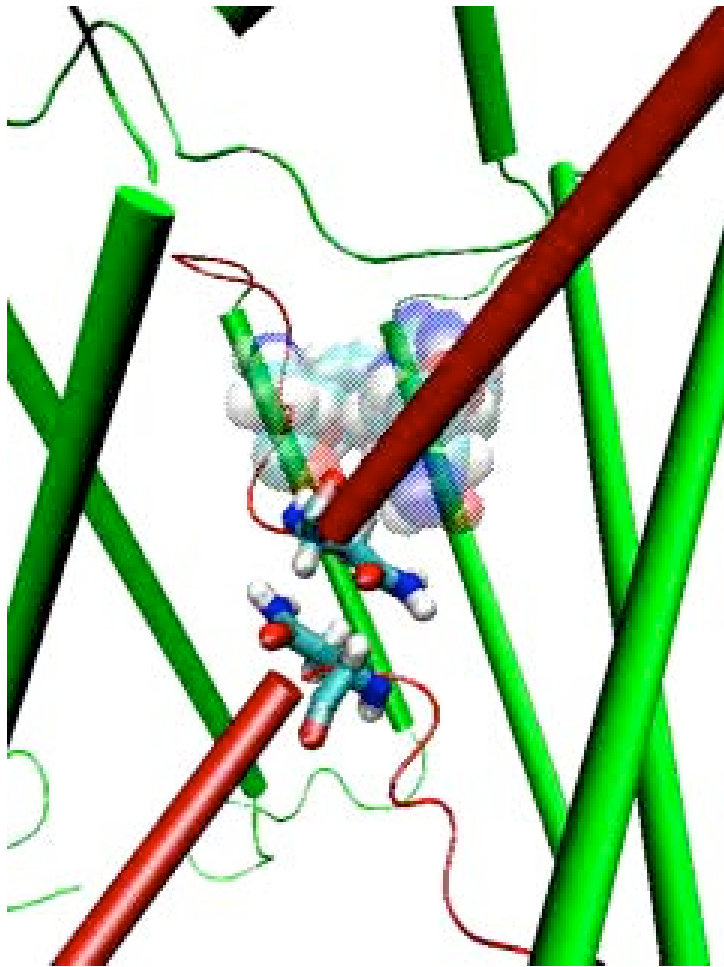
Confinement in Filter

- Selection occurs in most constrained region (induced fit)
- Selectivity probes shape, flexibility, hydrogen bonding.



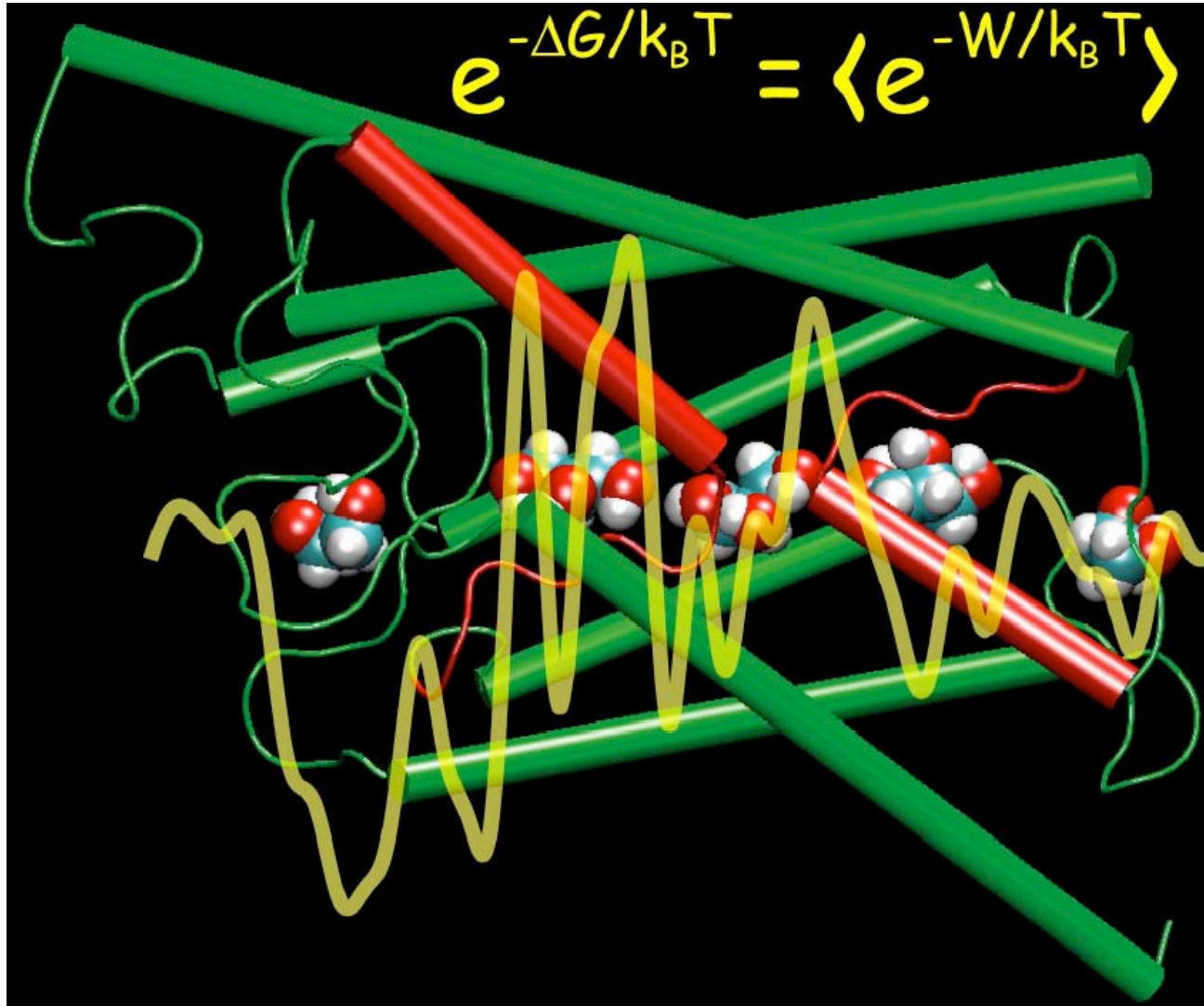
Results of Interactive Simulations

Restricted motion filter for all sugars



Dipole reversal at NPA

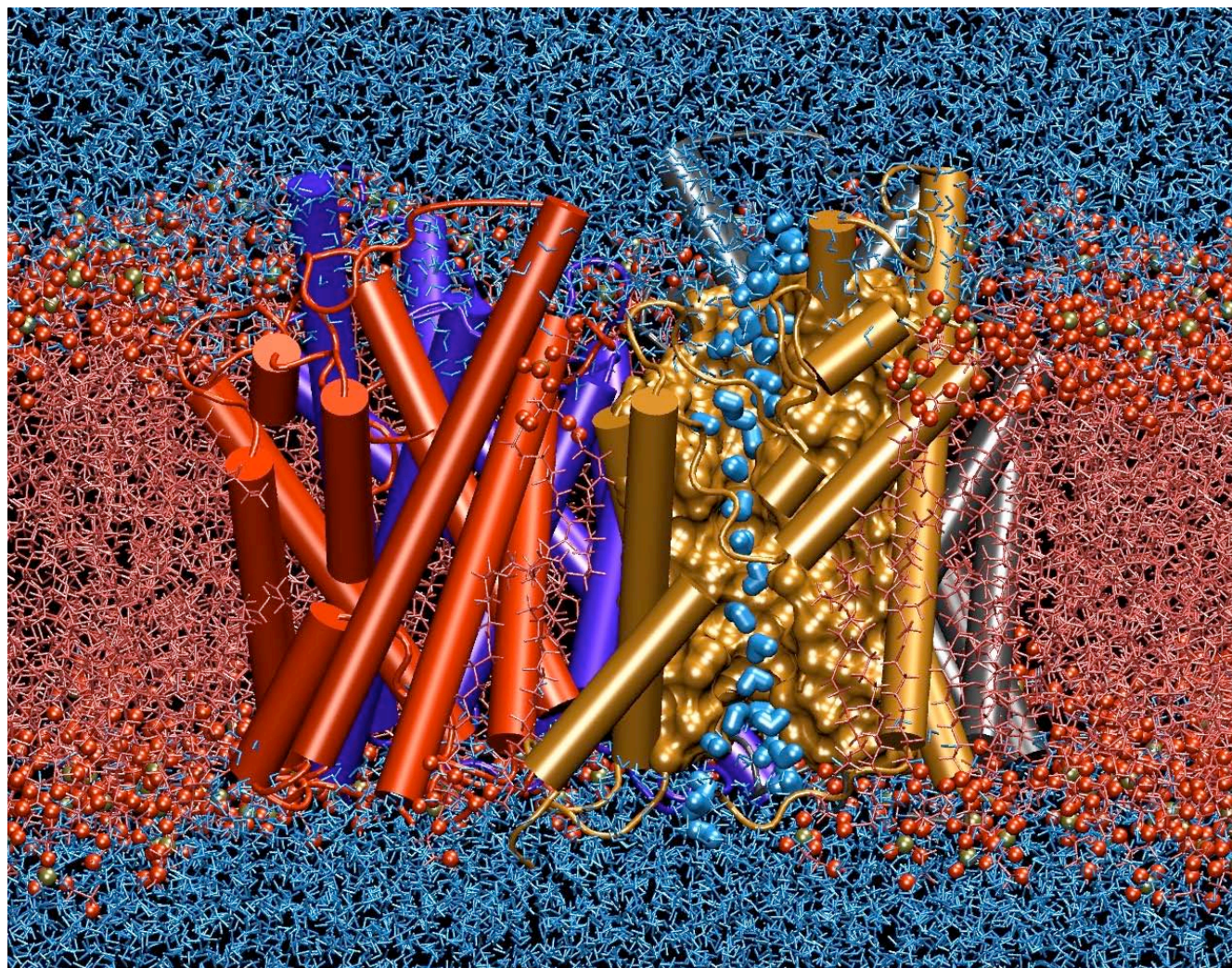
Resulting Potential of Mean Force



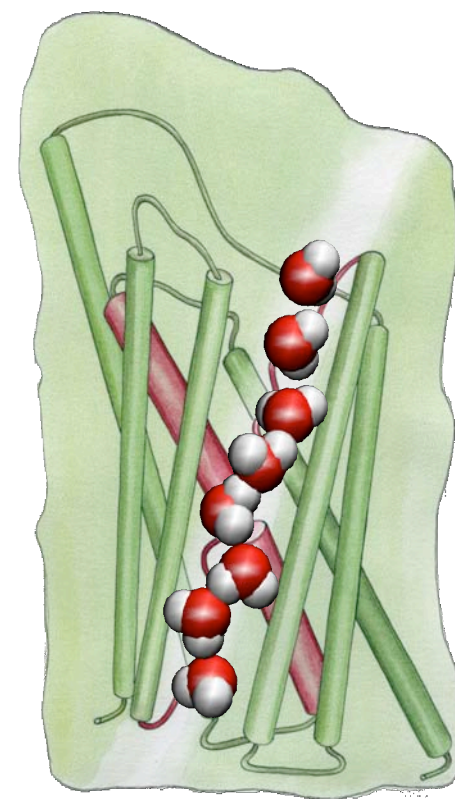
- The largest free energy barrier \approx **7.3 kcal/mol**
cf. Arrhenius activation energy measured:
9.6 \pm 1.5 kcal/mol, Borgnia and Agre (2001)

M. Jensen, S. Park, E.
Tajkhorshid, K. Schulten, *PNAS*
99:6731-6736 (2002)

Simulated System 3: GlpF With Only Water

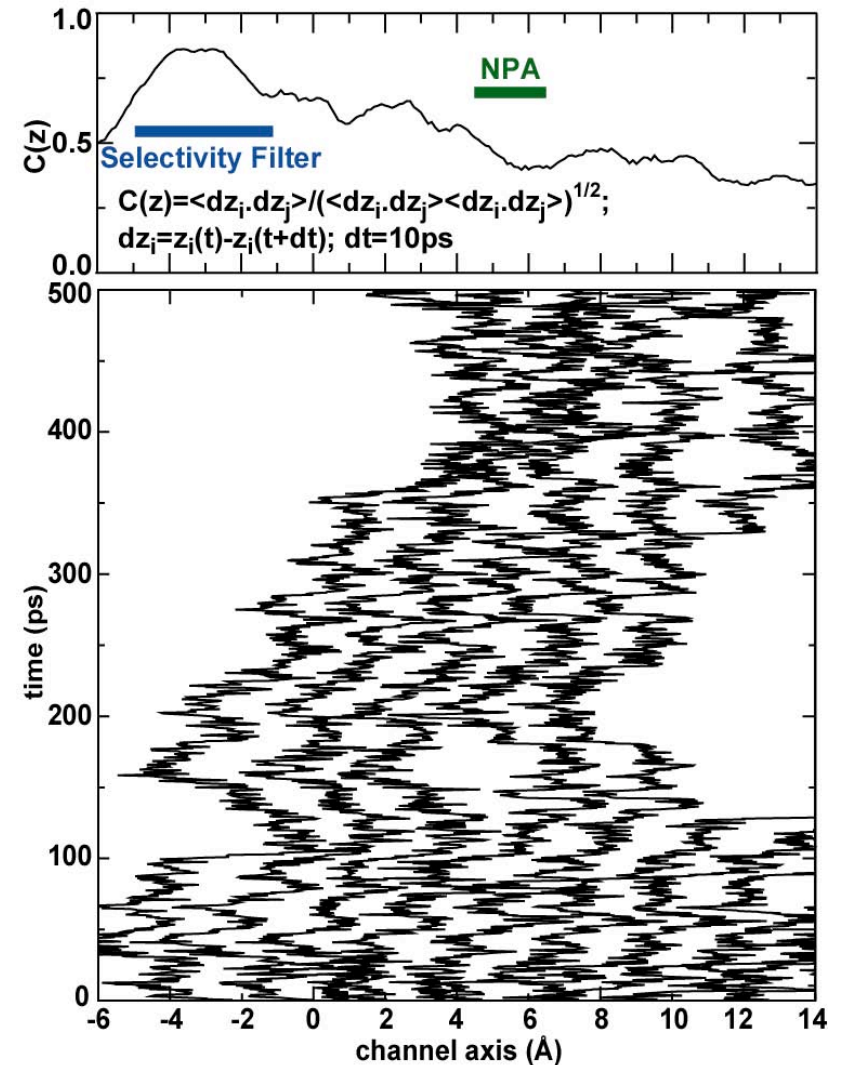
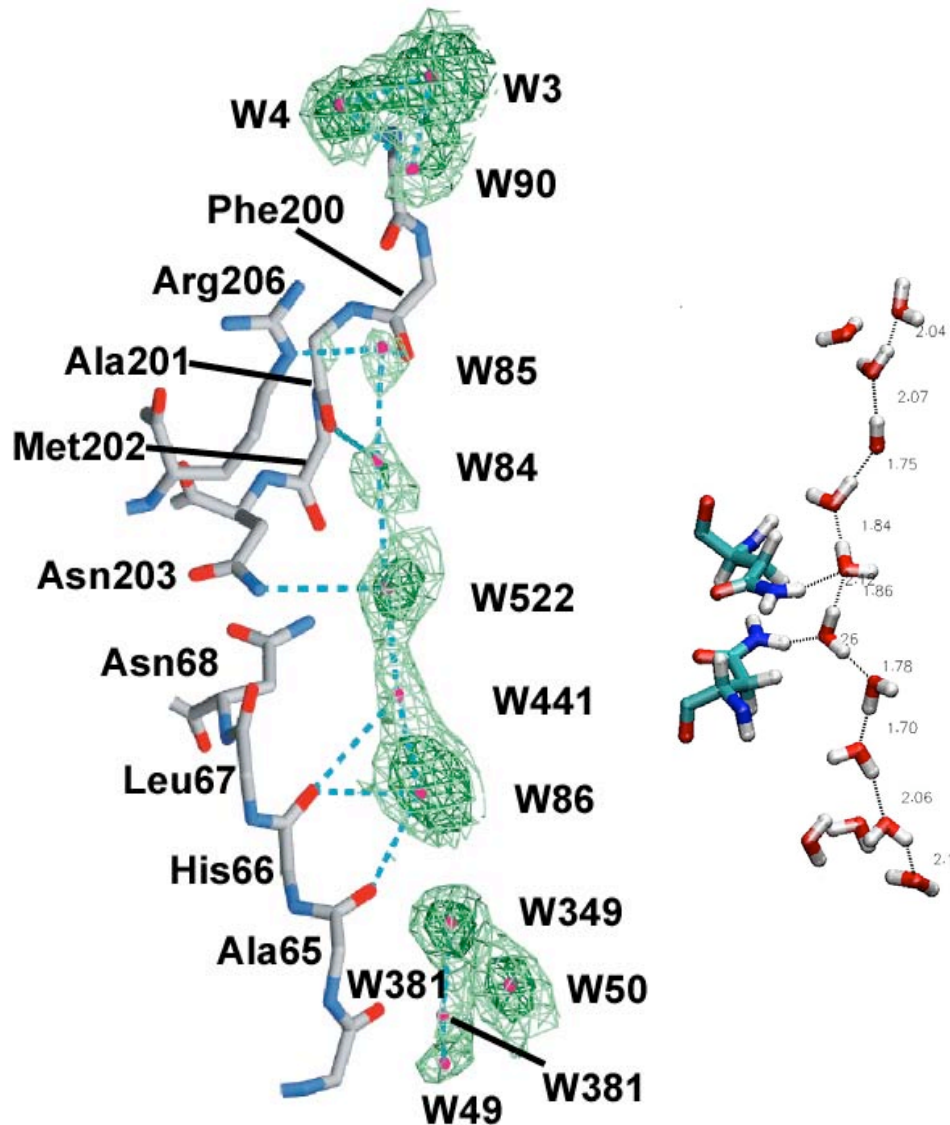


**5ns + 2ns
simulation**



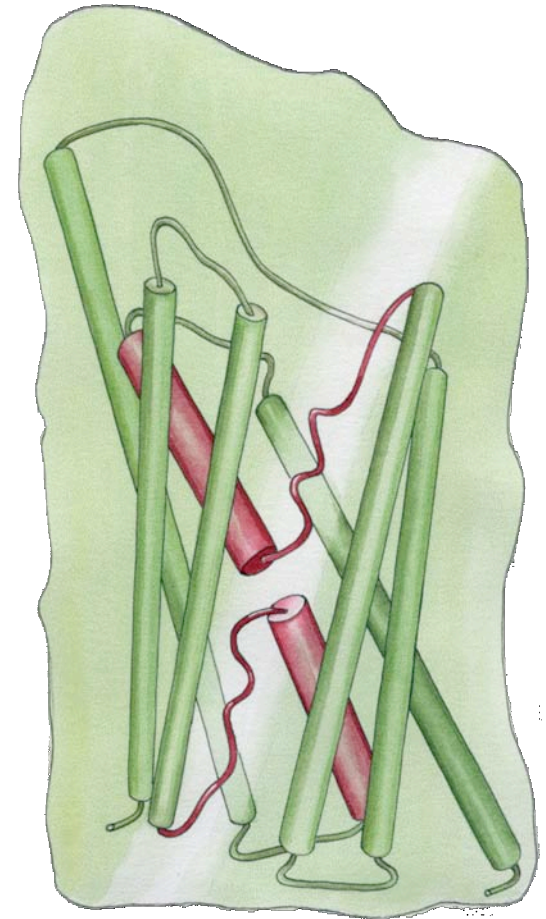
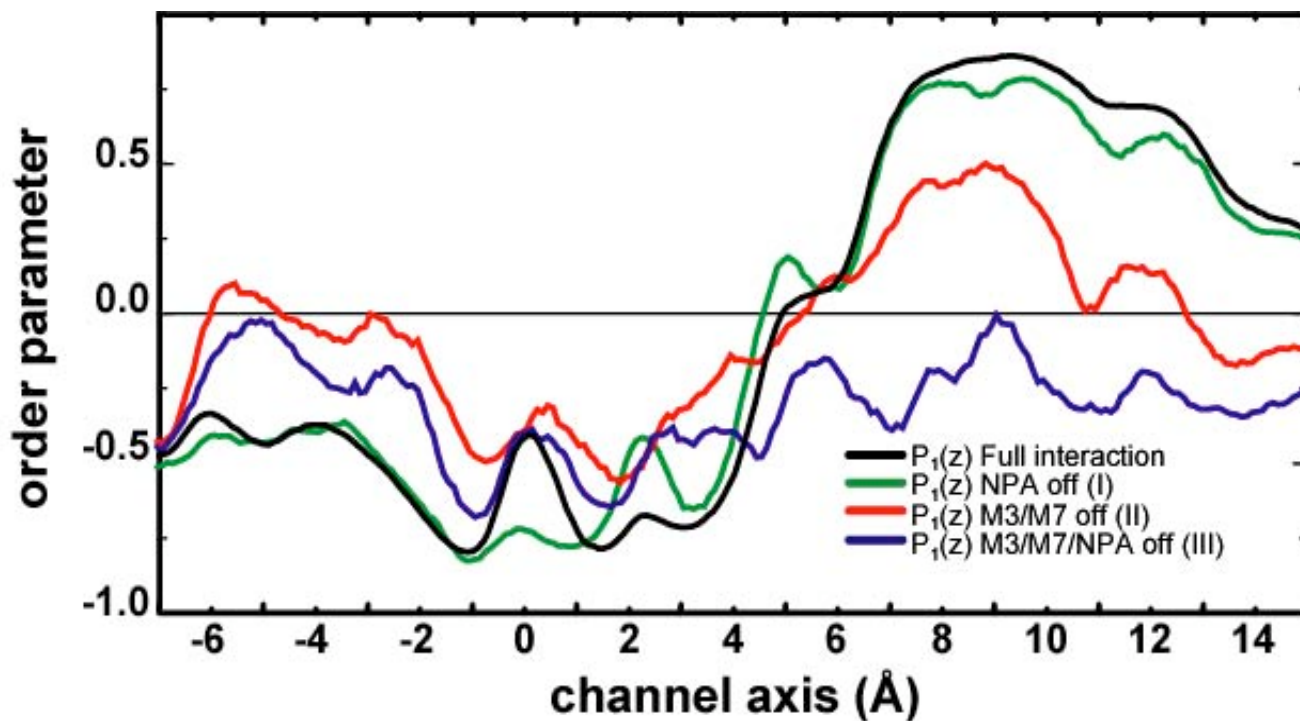
18 water molecules conducted / (4 monomers 4 ns) \rightarrow 1.125 water/monomer ns

Water Positions Determined by Simulations and Experiments



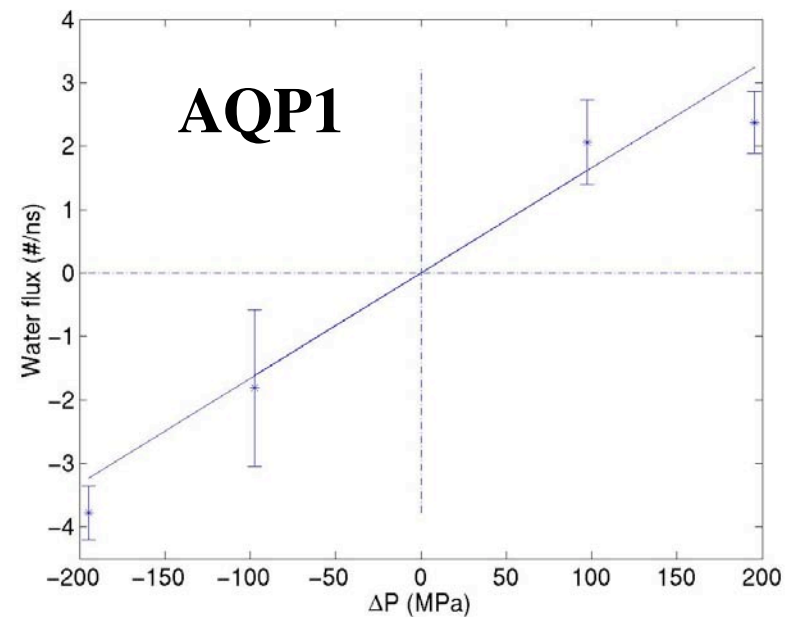
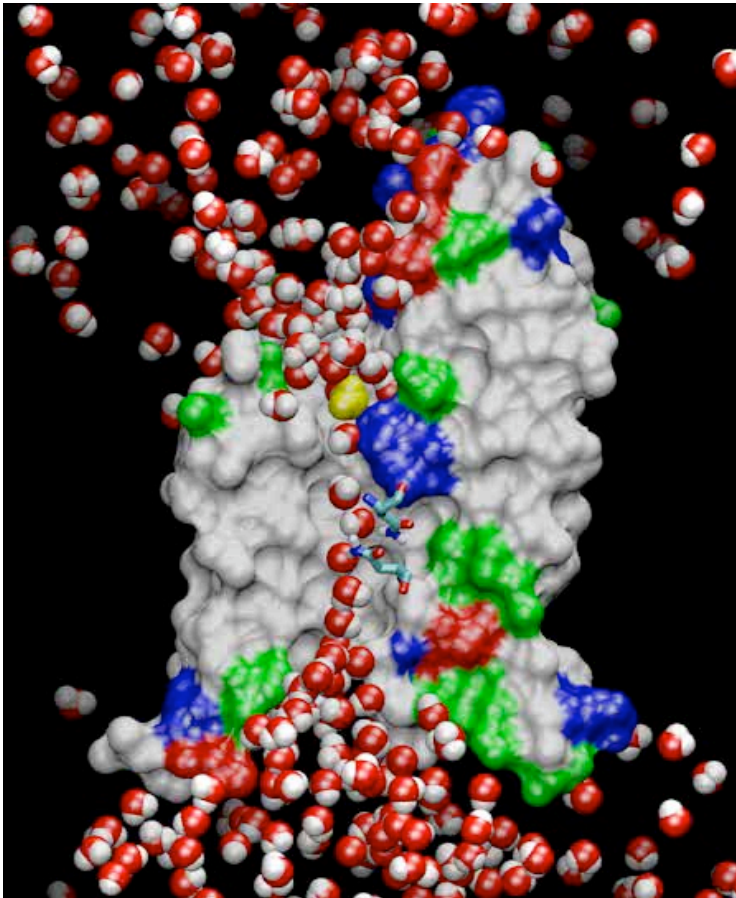
E. Tajkhorshid, P. Nollert, M. Jensen, L. J. W. Miercke, J. O'Connell, R. M. Stroud, and K. Schulten, *Science* 296, 525-530 (2002)

Electrostatic Stabilization of Water Bipolar Arrangement



E. Tajkhorshid, P. Nollert, M. Jensen, L. J. W. Miercke, J. O'Connell, R. M. Stroud, and K. Schulten, *Science* 296, 525-530 (2002)

Water flux vs. pressure difference

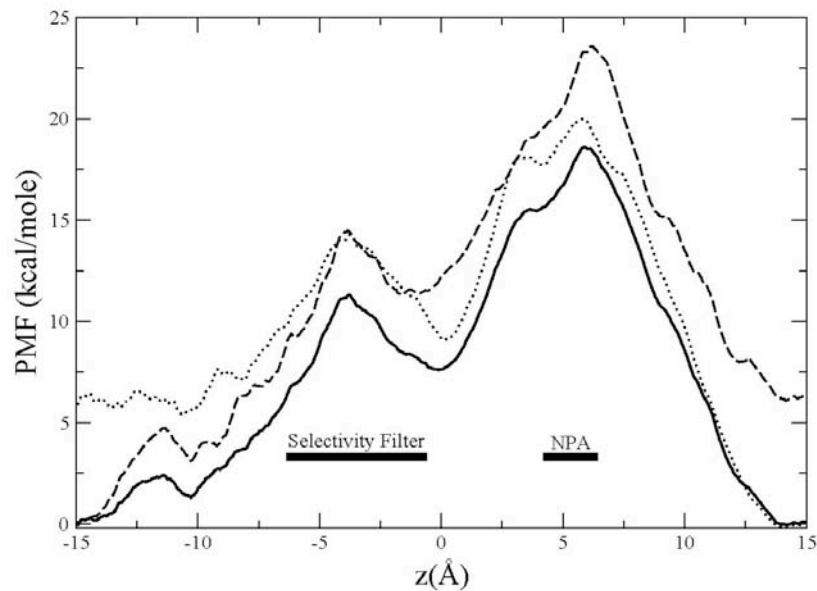


- Calculated p_f $(7.1 \pm 0.9) \times 10^{-14}$ cm³/s
- Experimental p_f values: $5 \sim 11 \times 10^{-14}$ cm³/s

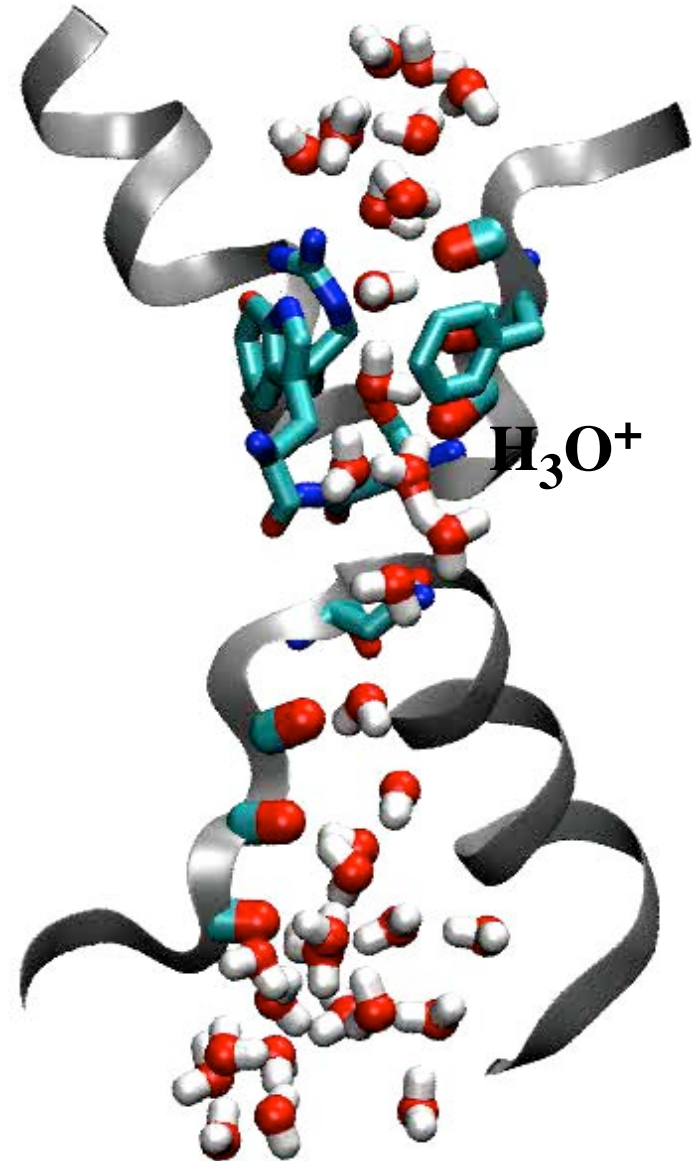
Proton Exclusion in Aquaporin Channel

Initial condition: H_3O^+ in center

Energy barrier preventing proton conduction



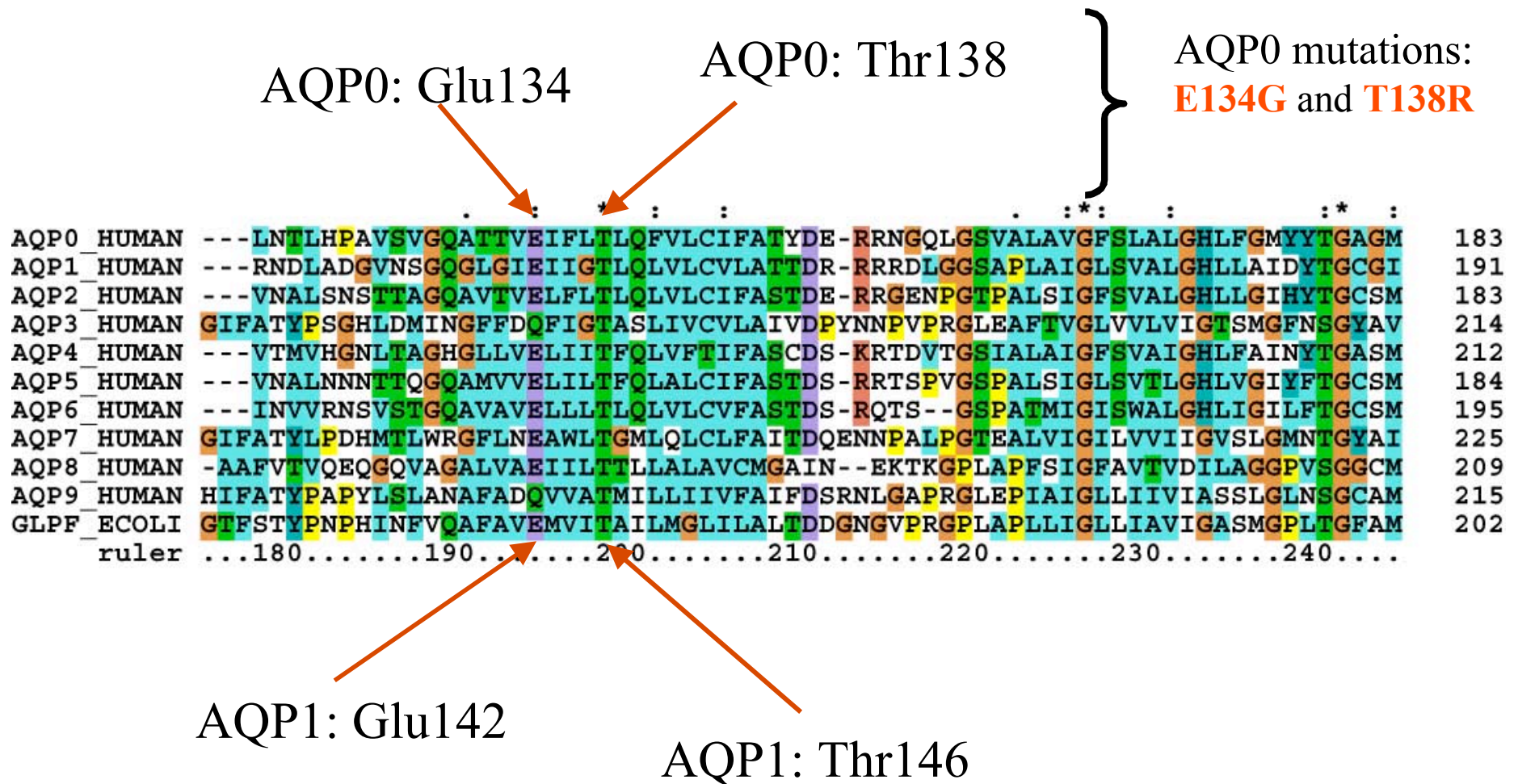
B. Ilan, E. Tajkhorshid, K. Schulten, G. Voth, *PROTEINS* 55: 223-228 (2004)



M. Hoffmann, E. Tajkhorshid, and K. Schulten (unpublished)

Genetically Inherited Cataracts

Impaired protein trafficking is suggested to be the main effect of these mutants, however, an impaired channel activity can also be involved.



E. Tajkhorshid et al (unpublished)

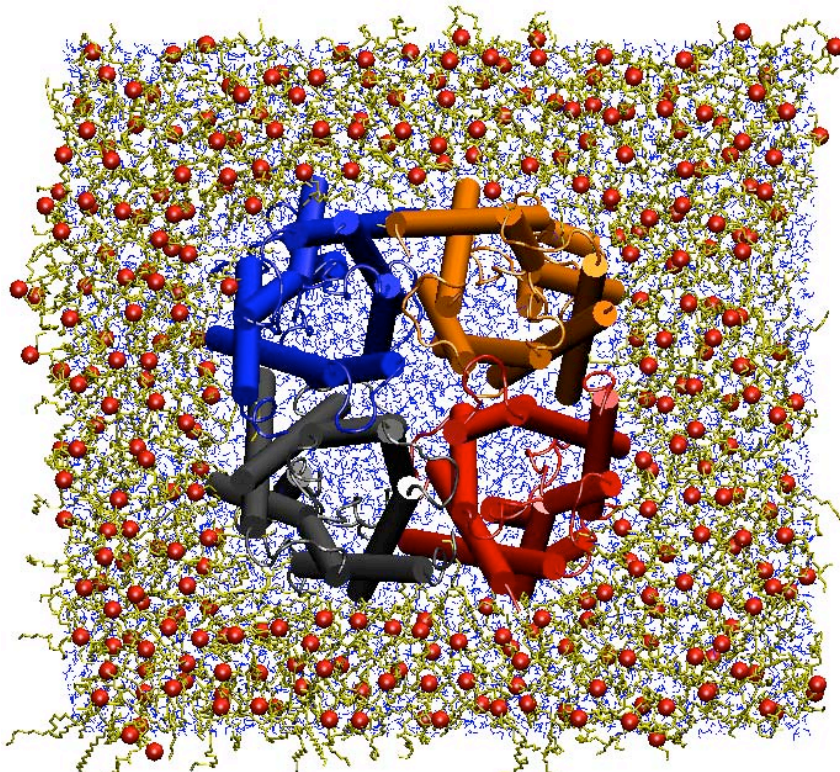
Probing Mutant Through Molecular Dynamics Simulations

Protein: ~ 15,000 atoms

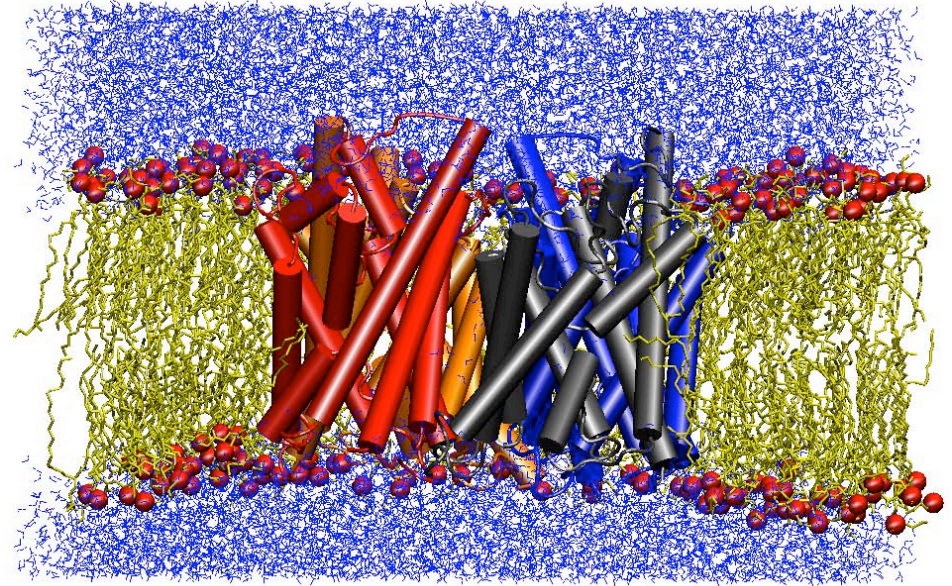
Lipids (POPE): ~ 40,000 atoms

Water: ~ 51,000 atoms

Total: ~ 106,000 atoms



E. Tajkhorshid et al (unpublished)



NAMD, CHARMM27, PME

NpT ensemble at 310 K

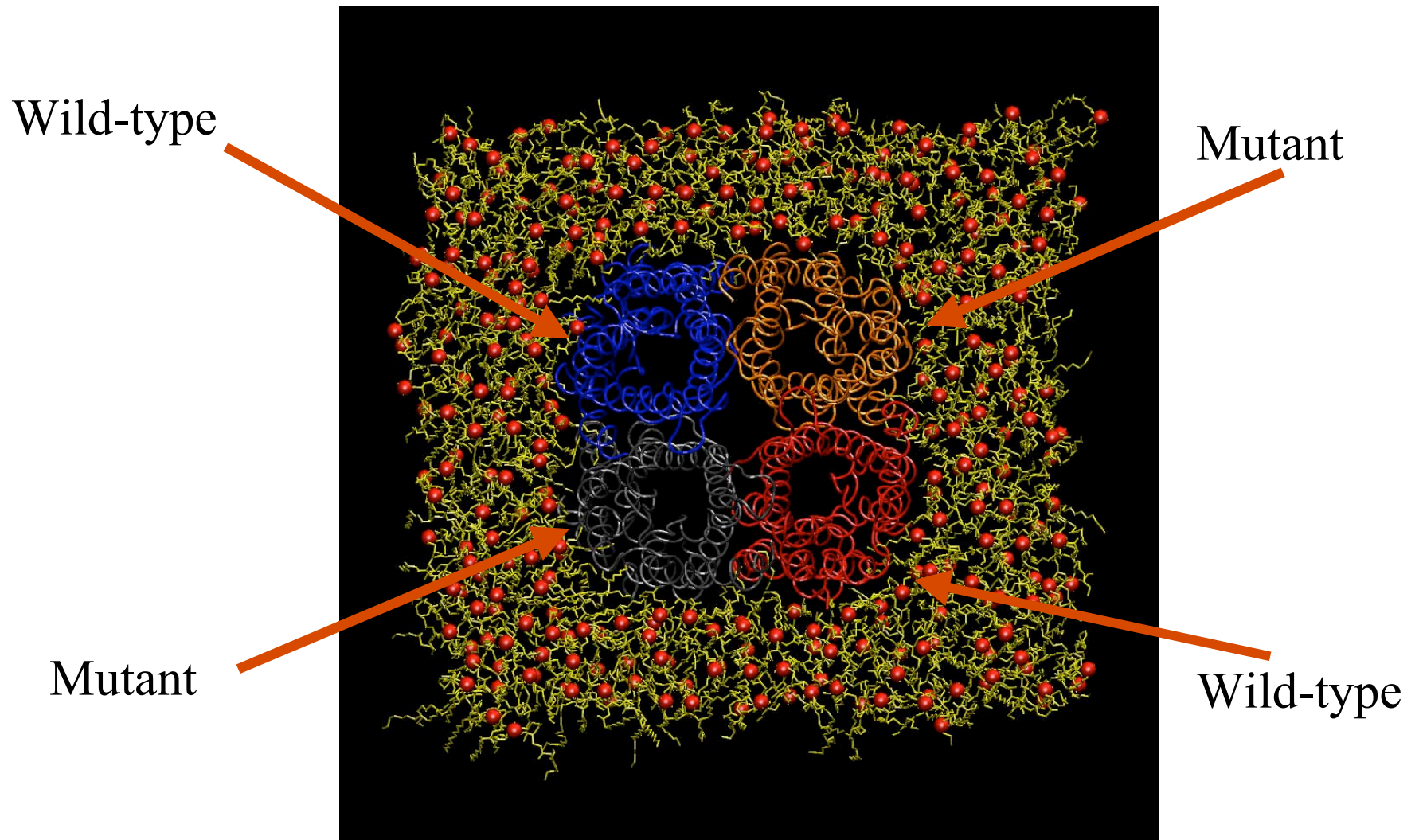
5 ns run of wild-type protein

15 ns simulation after mutation

10 days /ns – 32-proc Linux cluster

3.5 days/ns - 128 O2000 CPUs

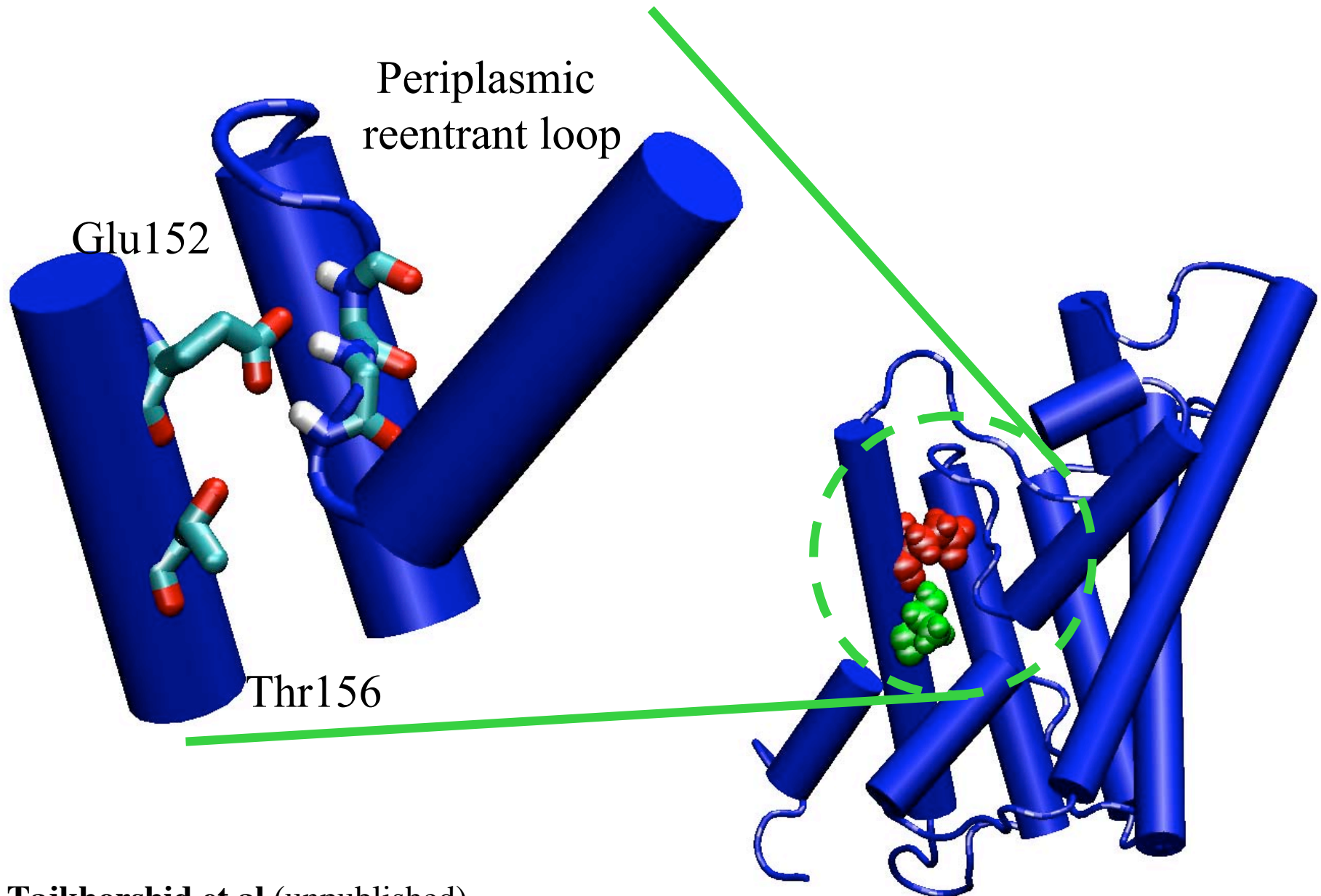
Point Mutations in the Tetramer



E152G in two diagonal monomers;
the other monomers were kept intact

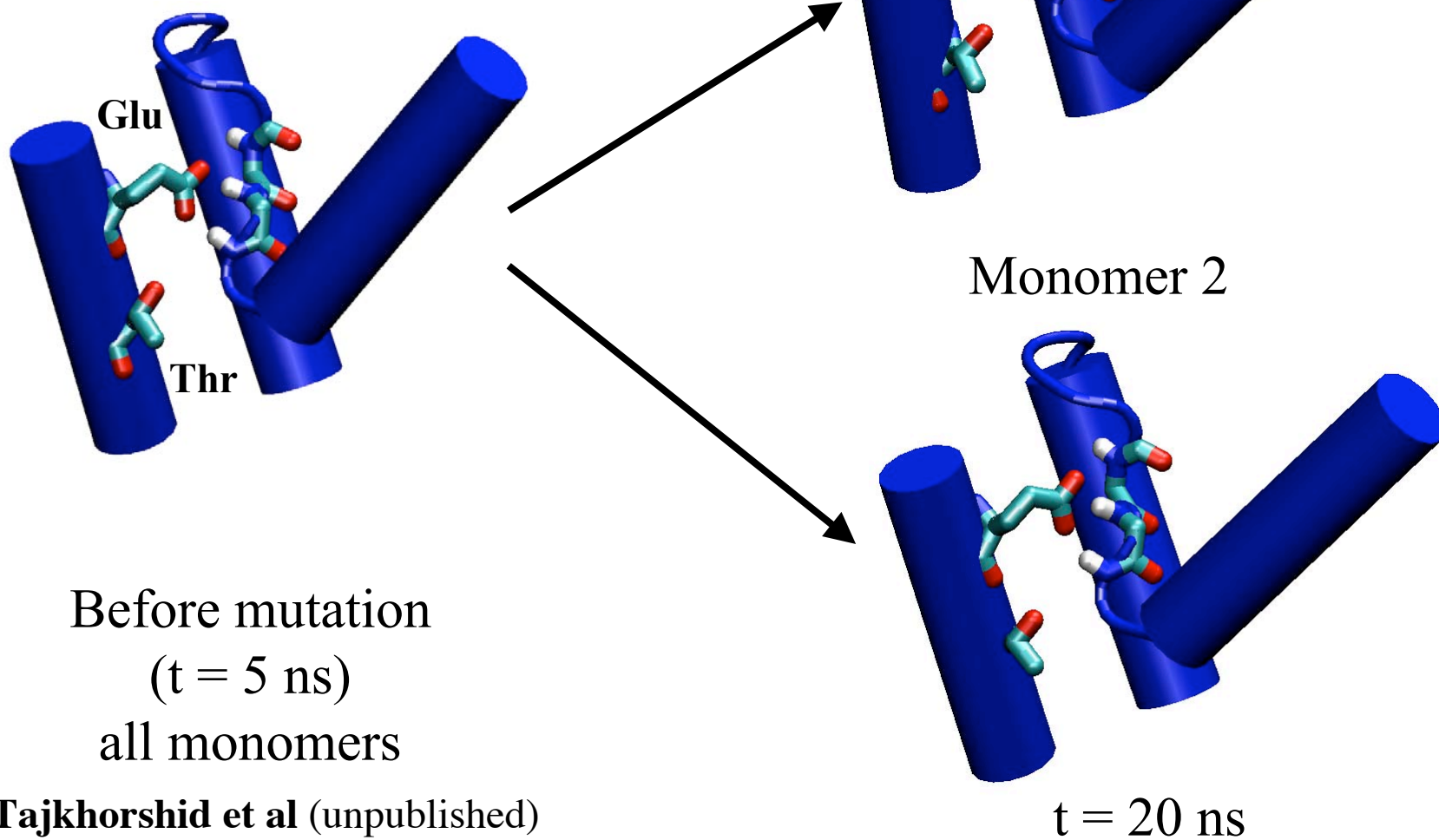
E. Tajkhorshid et al (unpublished)

Behaviour of Wild Type Aquaporin



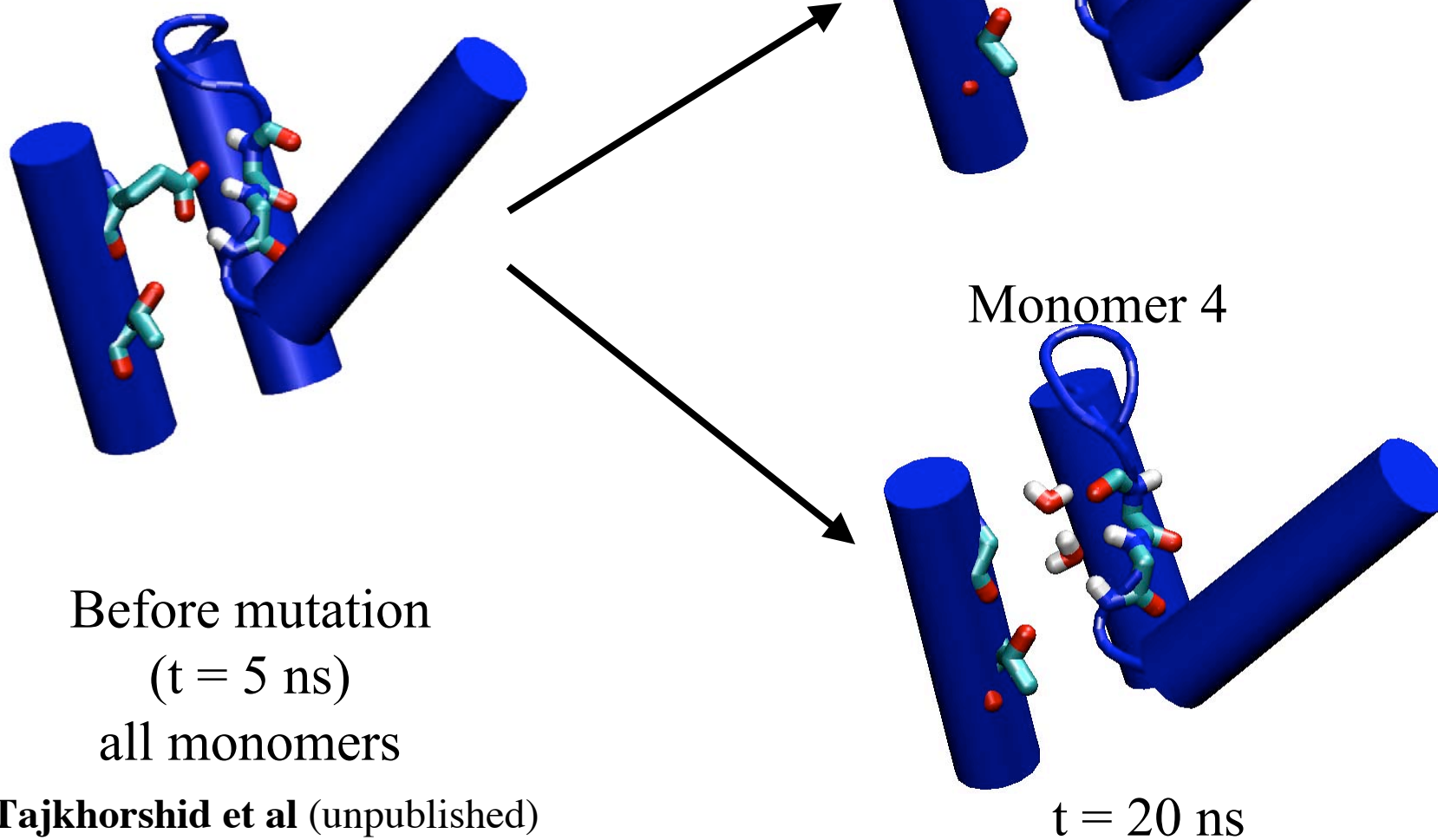
E. Tajkhorshid et al (unpublished)

Wildtype Conformational Changes



E. Tajkhorshid et al (unpublished)

Mutant Conformational Changes

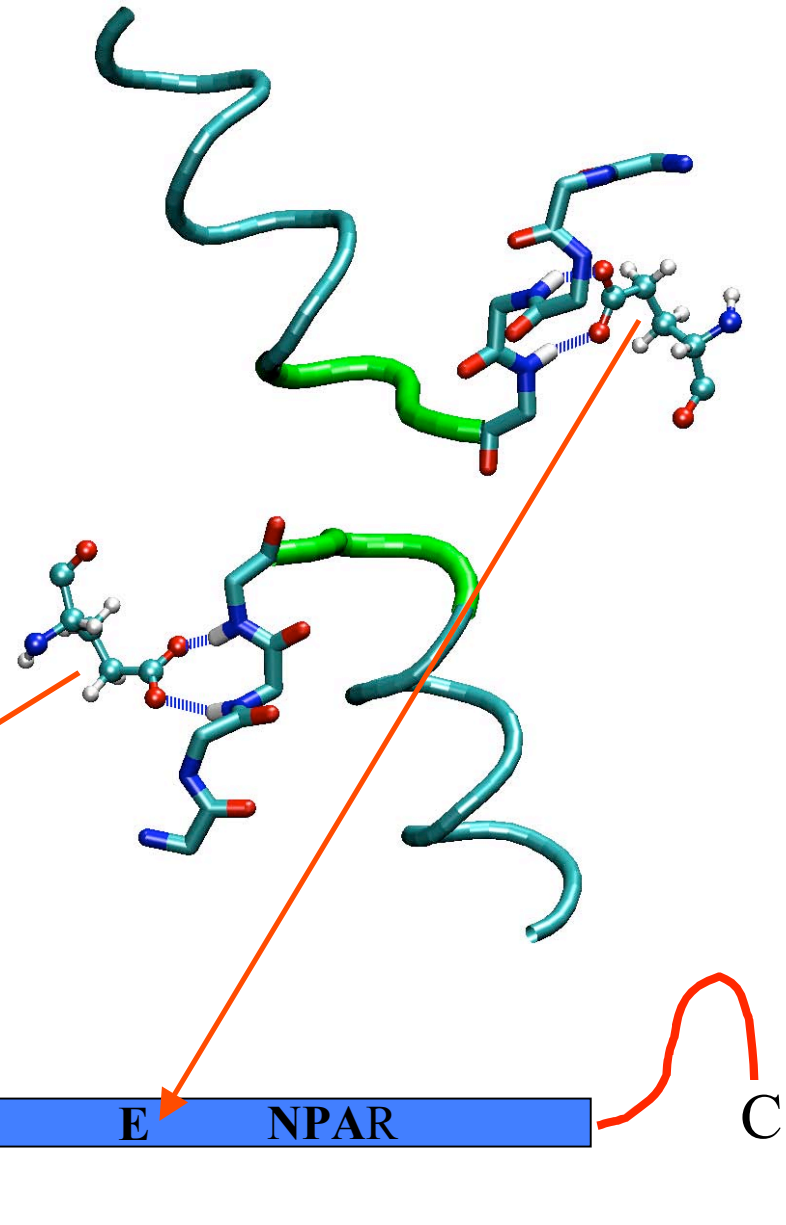


E. Tajkhorshid et al (unpublished)

Stability of the non-helical parts

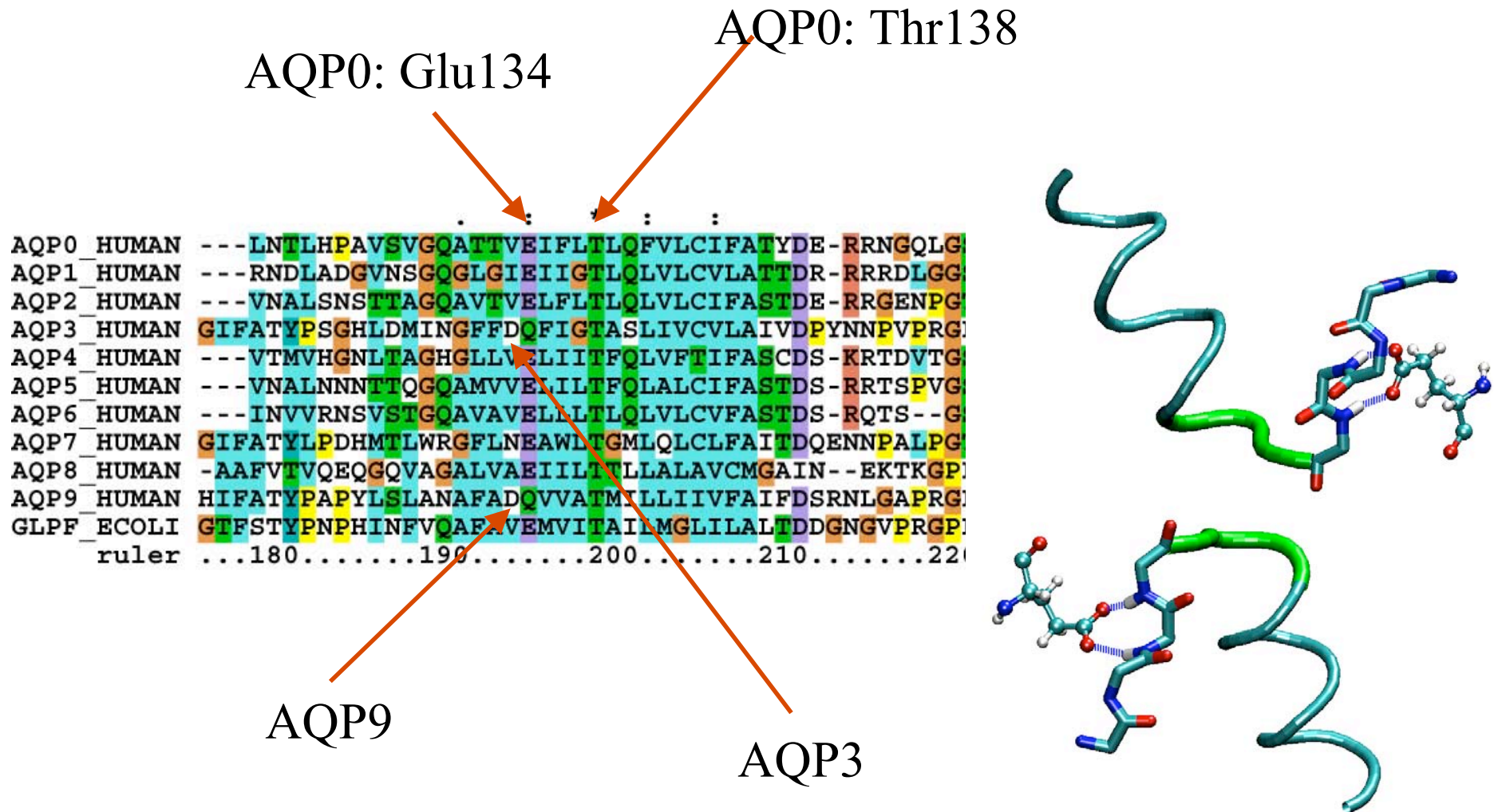
The **glutamate** residues stabilize the secondary structure of the inverted helices

The only glutamate residues buried in the transmembrane region of the channel



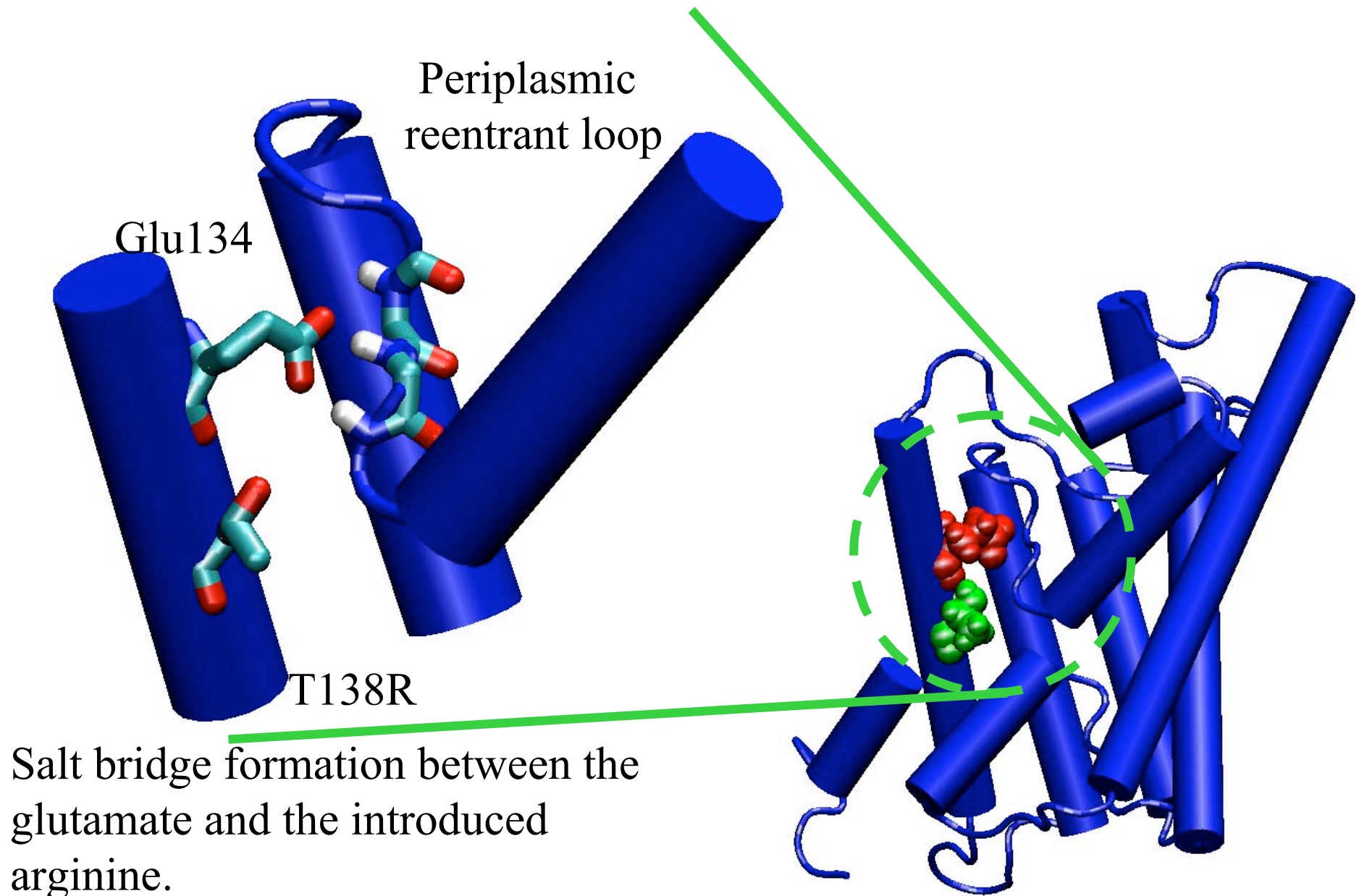
E. Tajkhorshid et al (unpublished)

Behaviour of Wild Type Aquaporins



Whenever GLU is missing, an ASP is present

Plausible Mechanism for T138R mutation

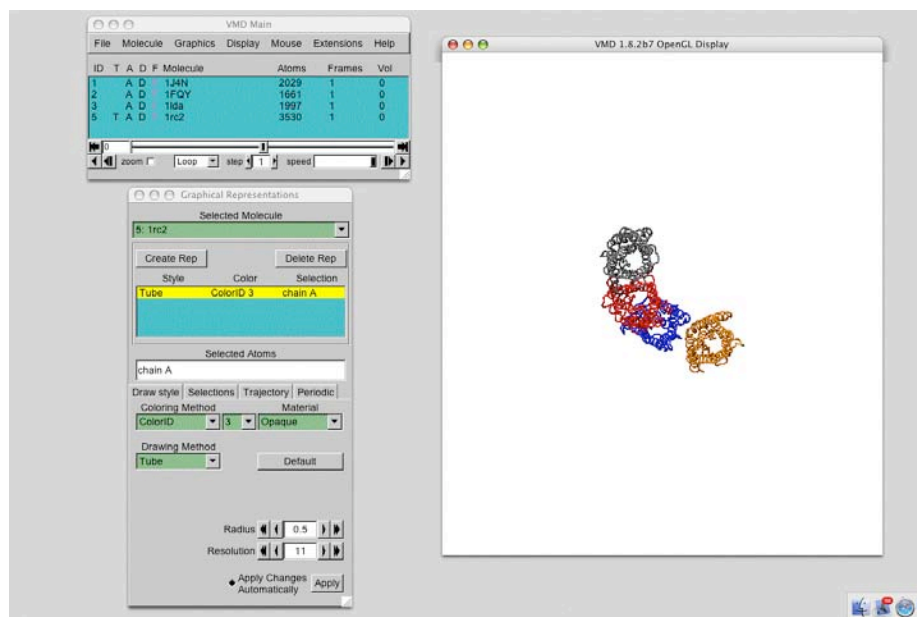


Tutorial

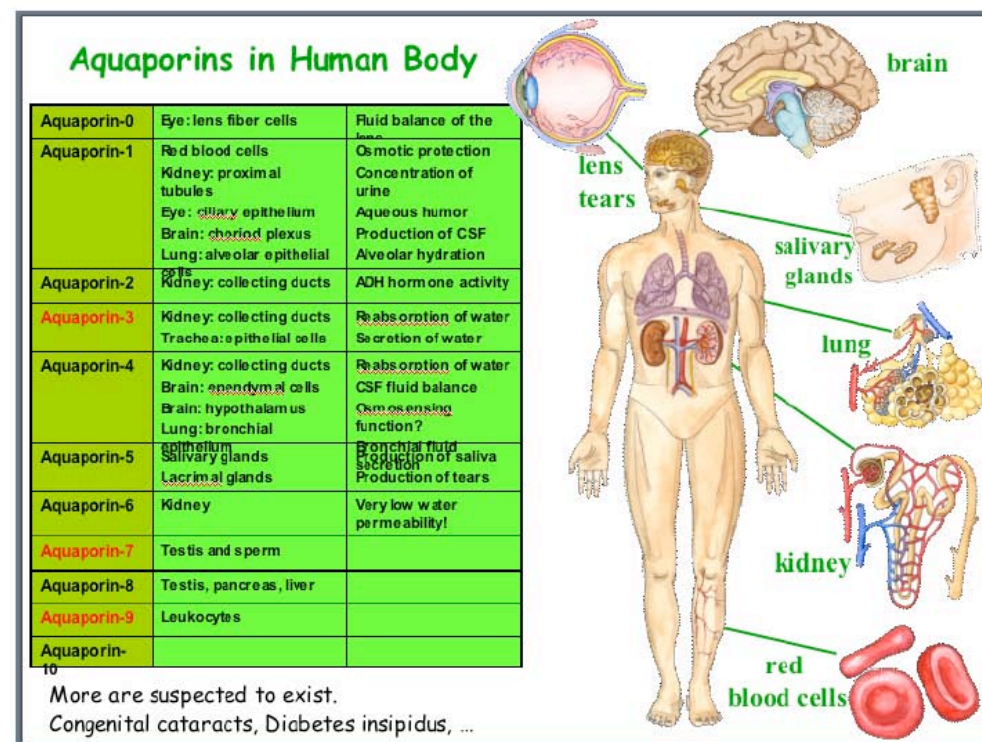
will be available at

www.ks.uiuc.edu

after release of VMD 1.8.3
 July 2004



Aquaporins



VMD Developers:

John Stone

Dan Wright

John Eargle

Fatemeh Khalili

Elizabeth Villa

Emad Tajkhorshid

Brijeet Dhaliwal

Zan Luthey-Schulten

Acknowledgments

Special thanks to
Emad Tajkhorshid

Zan Luthey-Schulten
(multiple sequence VMD)
Fatemeh Khalili-Araghi
Elizabeth Villa
(VMD tutorial)

UCSF Collaborators:
Peter Nollert
Robert Stroud
U. Paderborn Collaborator
Michael Hoffmann

**Theoretical Biophysics
Group, Beckman Institute,
UIUC**

**NIH Resource for
Macromolecular Modeling
and Bioinformatics**

