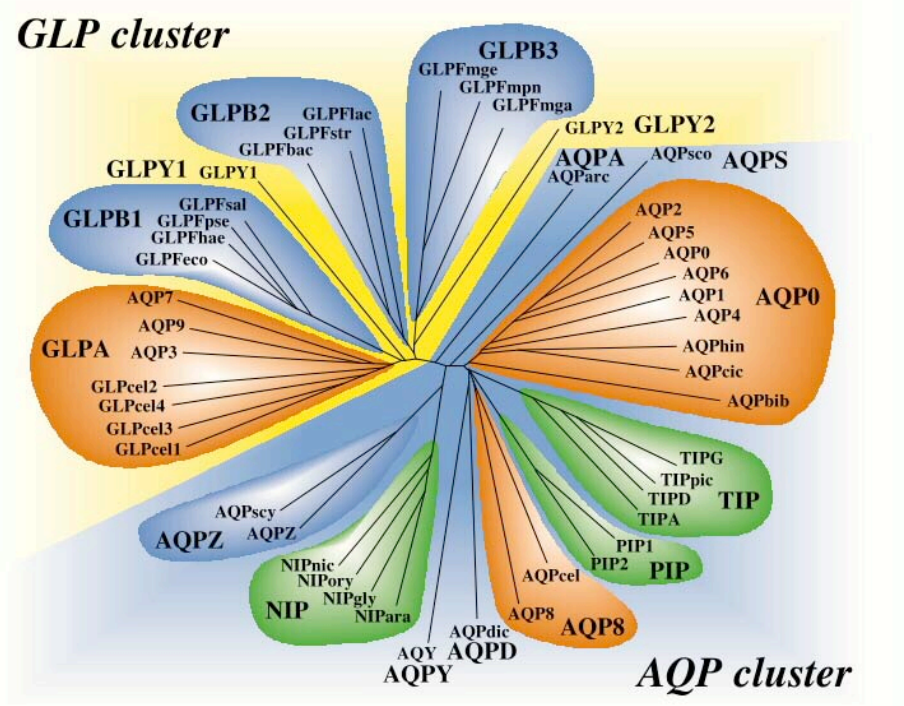
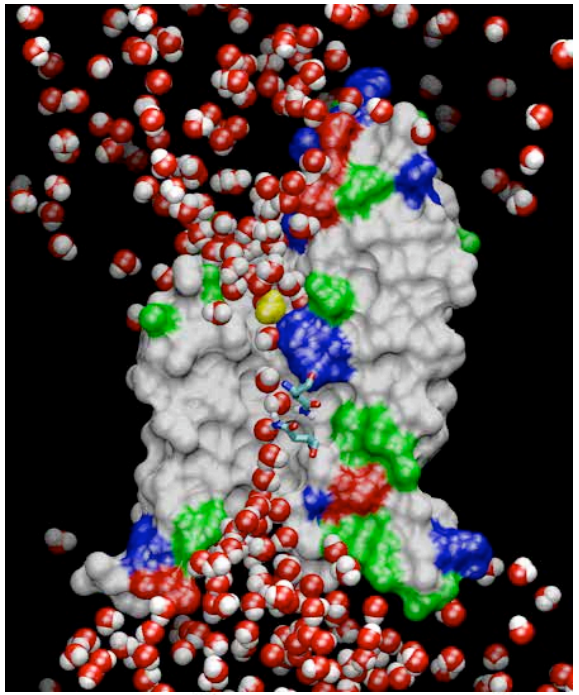
Physical Bioinformatics - A Case Study

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

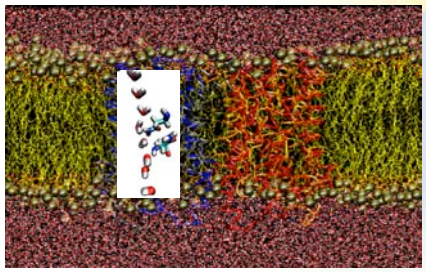
Klaus Schulten, U. Illinois at Urbana-Champaign





```
                                 .    :    *    :    :                         .    :*:    :          :*   :
AQP0_HUMAN  ---LNTLHPAVSVGQATTVEIFLTLQFVLCIFATYDE-RRNGQLGSVALAVGFSLALGHLFGMYYTGAGM   183
AQP1_HUMAN  ---RNDLADGVNSGQGLGIEIIGTLQLVLCVLATTDR-RRRDLGGSAPLAIGLSVALGHLLAIDYTGCGI   191
AQP2_HUMAN  ---VNALSNSTTAGQAVTVELFLTLQLVLCIFASTDE-RRGENPGTPALSIGFSVALGHLLGIHYTGCSM   183
AQP3_HUMAN  GIFATYPSGHLDMINGFFDQFIGTASLIVCVLAIVDPYNNPVPRGLEAFTVGLVVLVIGTSMGFNSGYAV   214
AQP4_HUMAN  ---VTMVHGNLTAGHGLLVELIITFQLVFTIFASCDS-KRTDVTGSIALAIGFSVAIGHLFAINYTGASM   212
AQP5_HUMAN  ---VNALNNNTTQGQAMVVELILTFQLALCIFASTDS-RRTSPVGSPALSIGLSVTLGHLVGIYFTGCSM   184
AQP6_HUMAN  ---INVVRNSVSTGQAVAVELLLTLQLVLCVFASTDS-RQTS--GSPATMIGISWALGHLIGILFTGCSM   195
AQP7_HUMAN  GIFATYLPDHMTLWRGFLNEAWLTGMLQLCLFAITDQENNPALPGTEALVIGILVVIIGVSLGMNTGYAI   225
AQP8_HUMAN  -AAFVTVQEQGQVAGALVAEIILTTLLALAVCMGAIN--EKTKGPLAPFSIGFAVTVDILAGGPVSGGCM   209
AQP9_HUMAN  HIFATYPAPYLSLANAFADQVVATMILLIIVFAIFDSRNLGAPRGLEPIAIGLLIIVIASSLGLNSGCAM   215
GLPF_ECOLI  GTFSTYPNPHINFVQAFAVEMVITAILMGLILALTDDGNGVPRGPLAPLLIGLLIAVIGASMGPLTGFAM   202
      ruler ...180.......190.......200.......210.......220.......230.......240....
```

The Aquaporin Superfamily

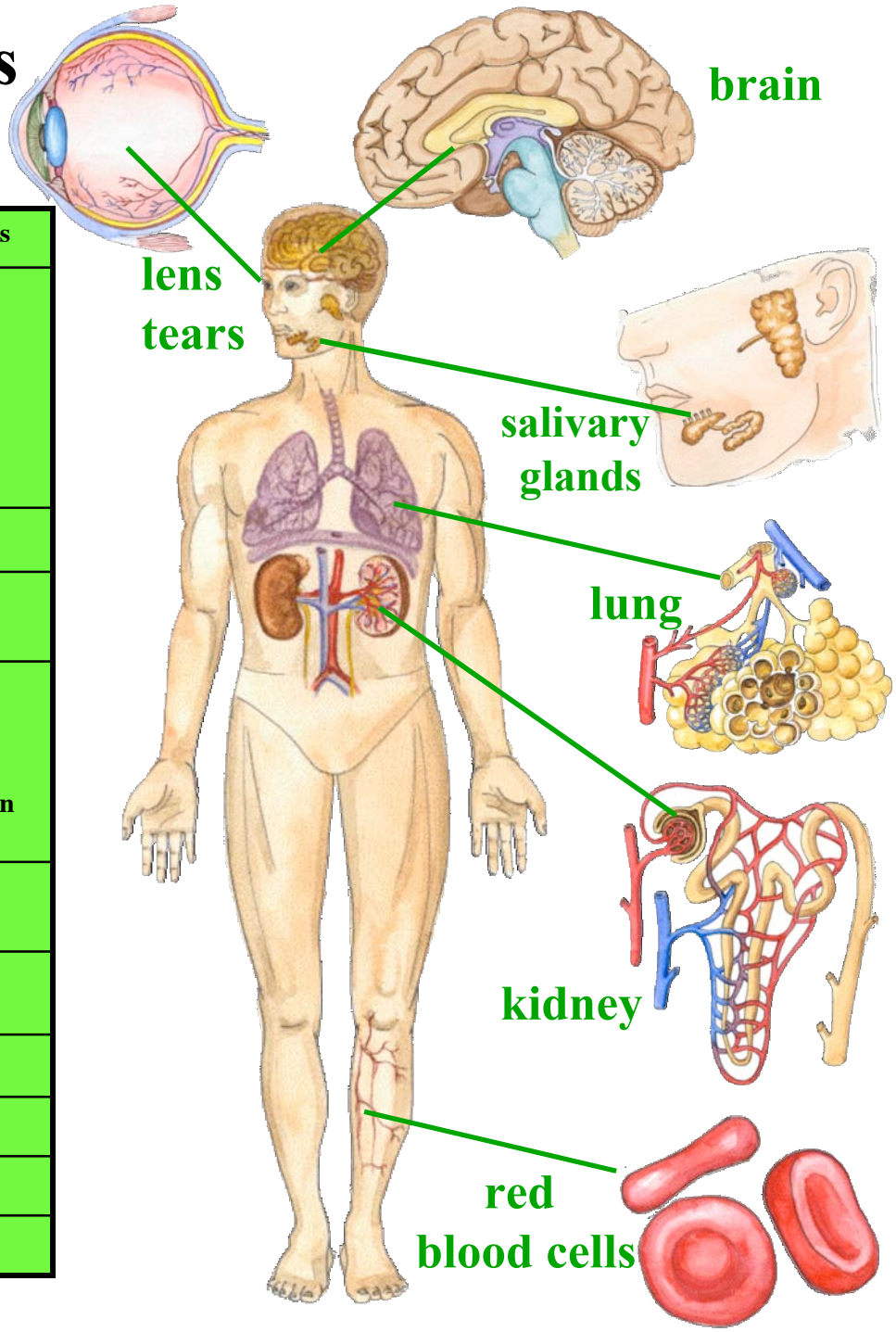


Heymann and Engel *News Physiol. Sci.* **14**, 187 (1999)

Water and Glycerol Channels in the Human Body

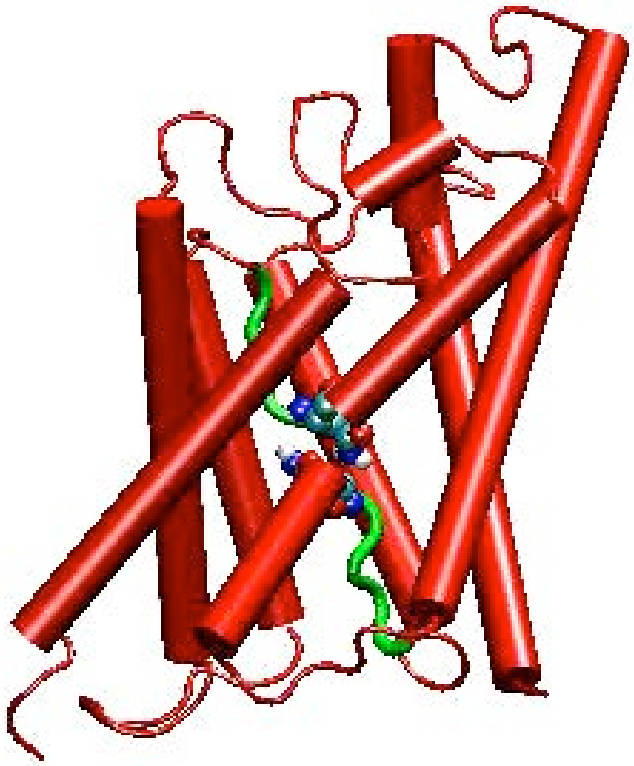
Aquaporin-0	Eye: lens fiber cells	Fluid balance of the lens
Aquaporin-1	Red blood cells Kidney: proximal tubules Eye: ciliary epithelium Brain: choriod plexus Lung: alveolar epithelial cells	Osmotic protection Concentration of urine Aqueous humor Production of CSF Alveolar hydration
Aquaporin-2	Kidney: 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

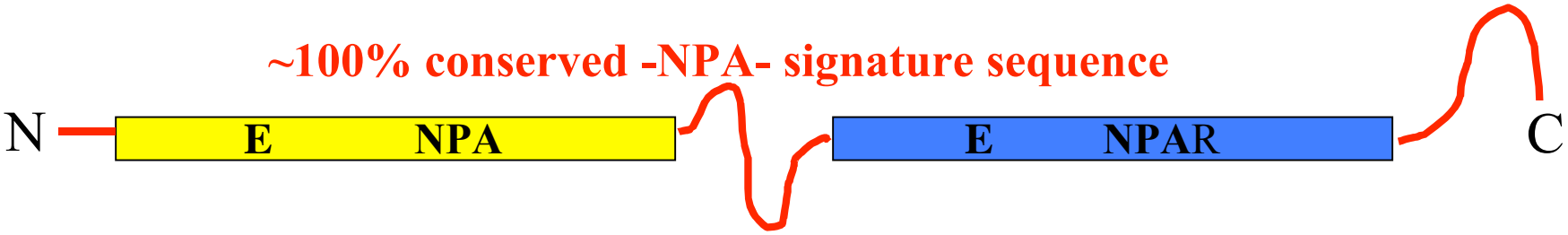
- Water and glycerol transport
- Exclusion of ions and protons
- Tetrameric arrangement in membrane



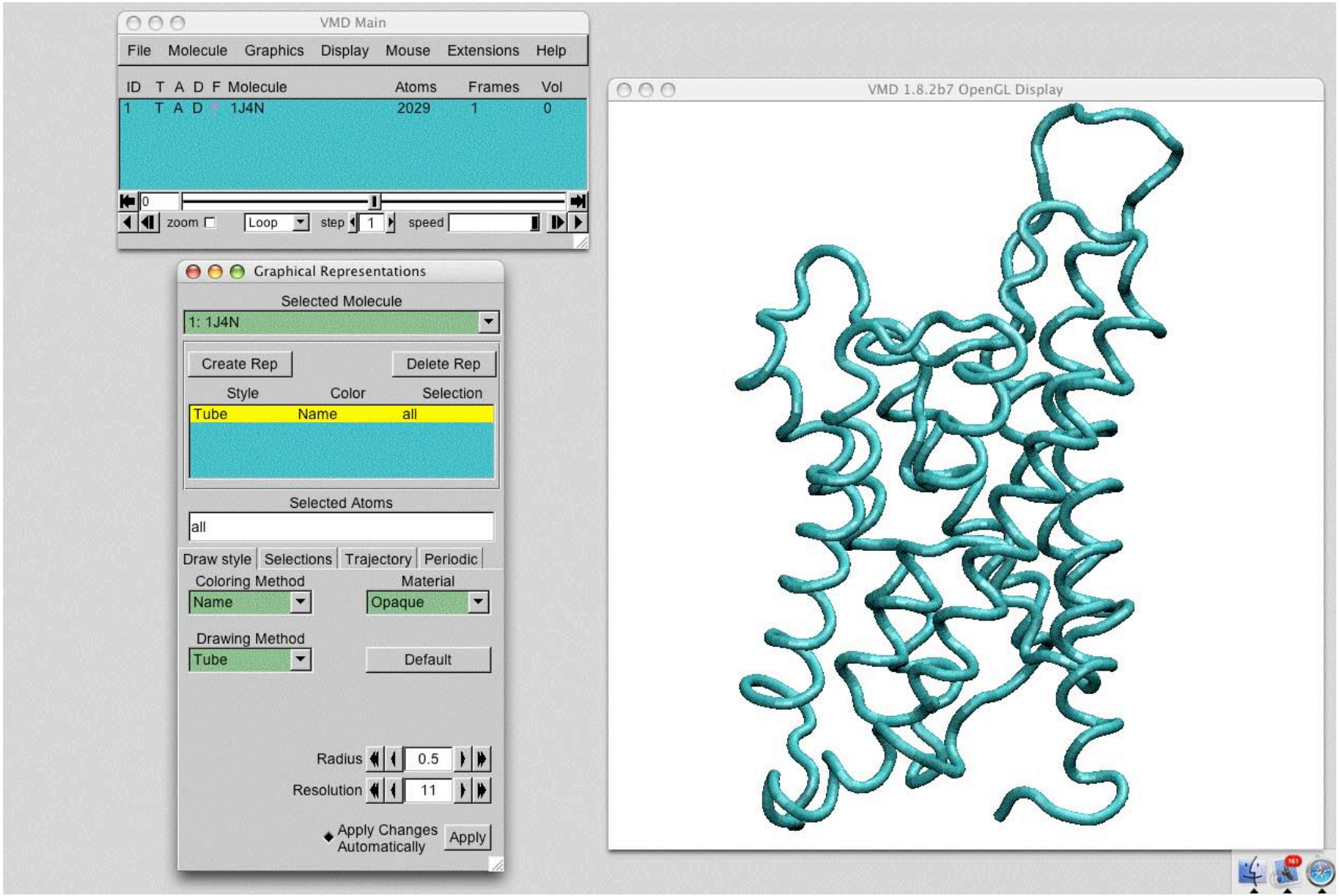
Aquaporins of known structure:

GlpF – E. coli glycerol channel (aquaglyceroporin) – Fu, et al., Science (2000)

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

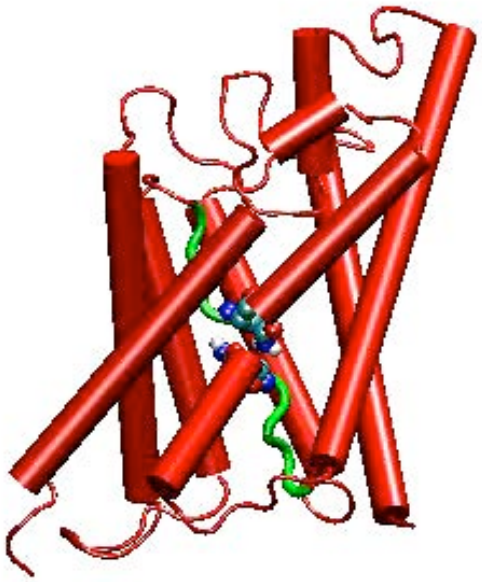


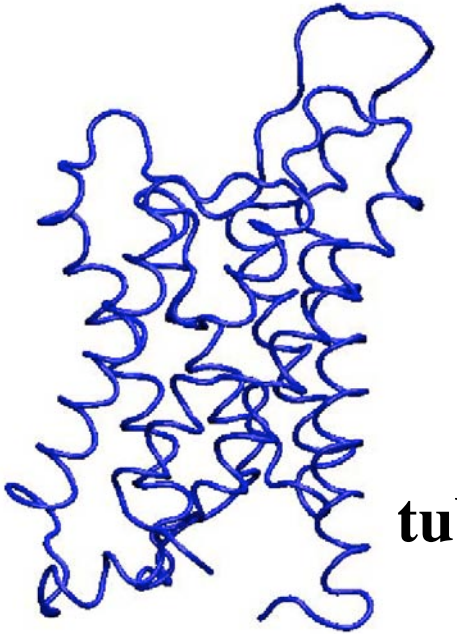
Load Aquaporin 1J4N into VMD



VMD Permits Different Rendering Styles

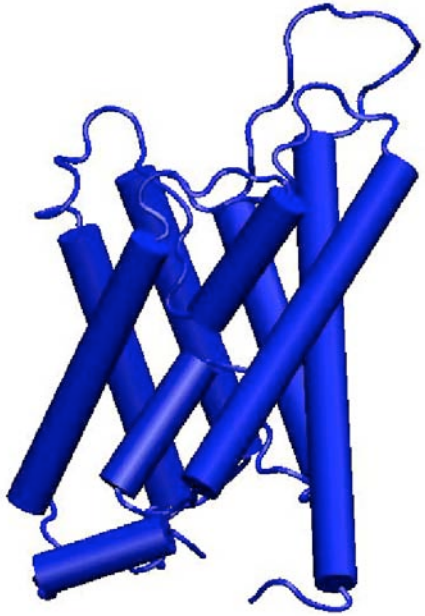
movie

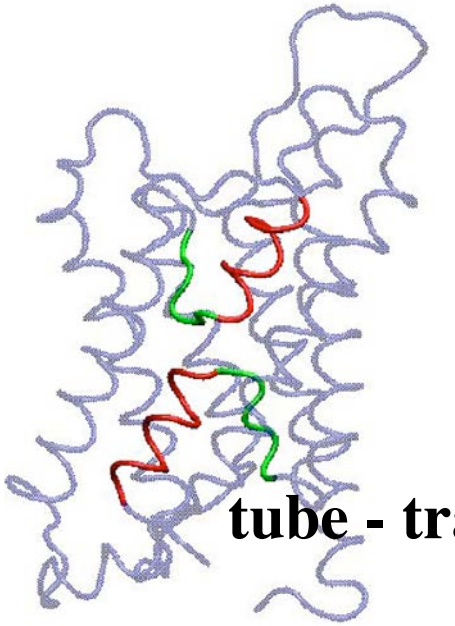




tube

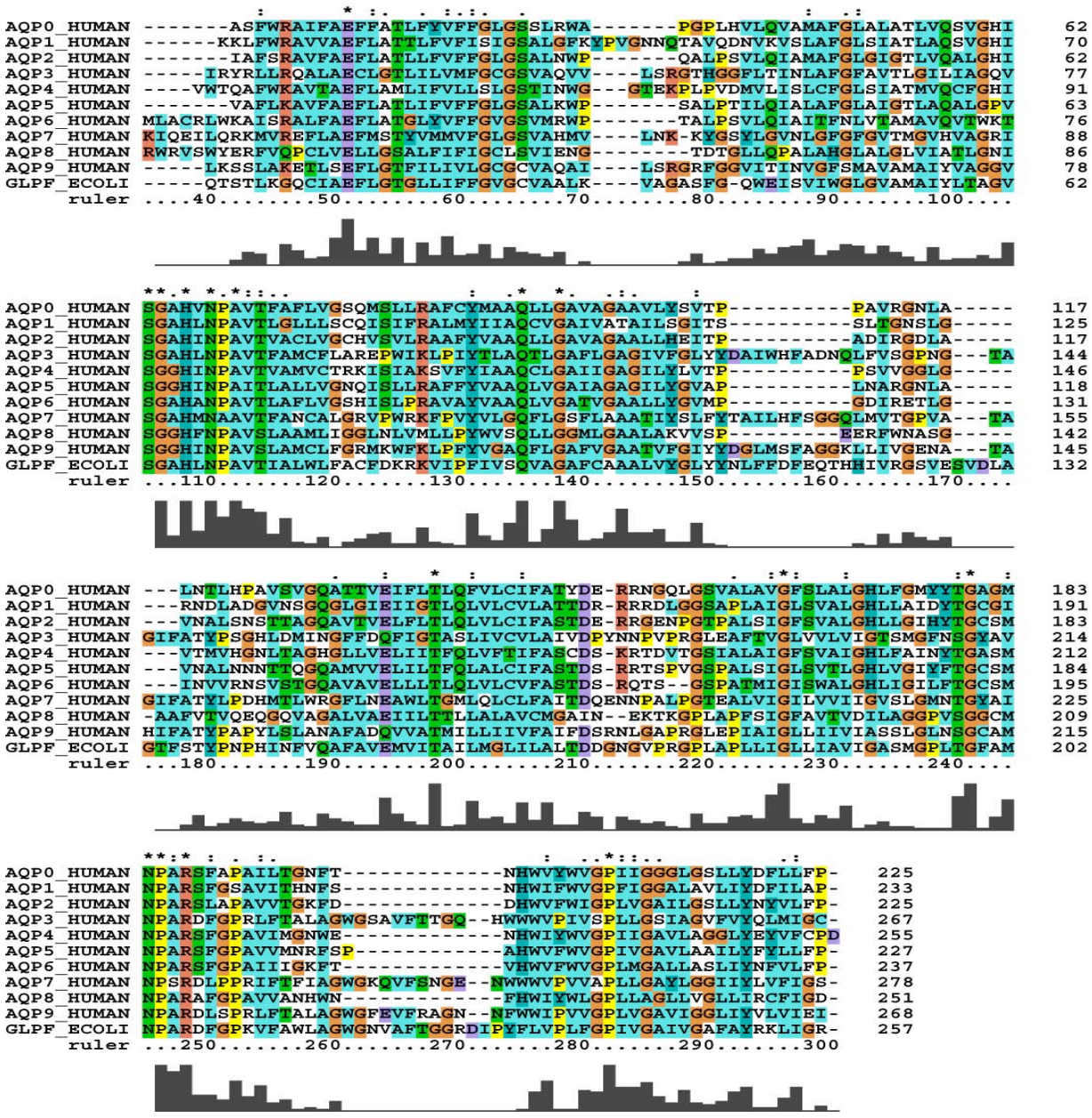
cartoon



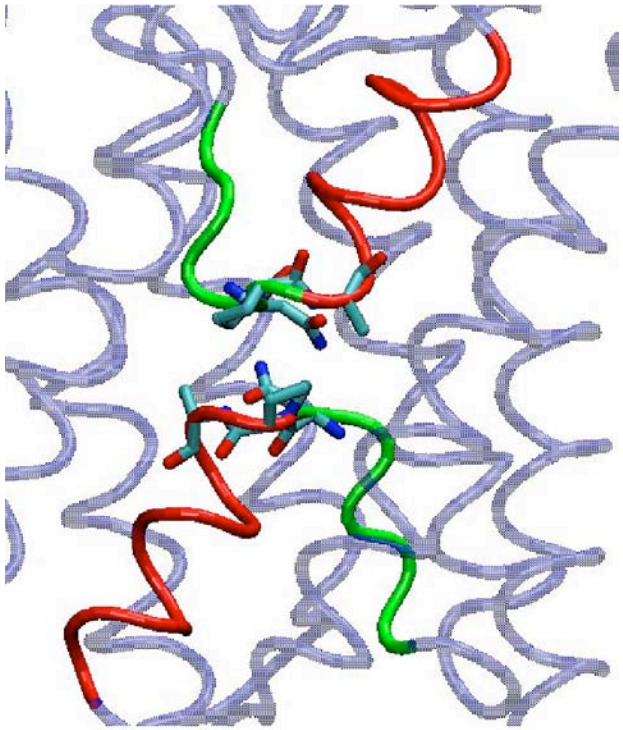


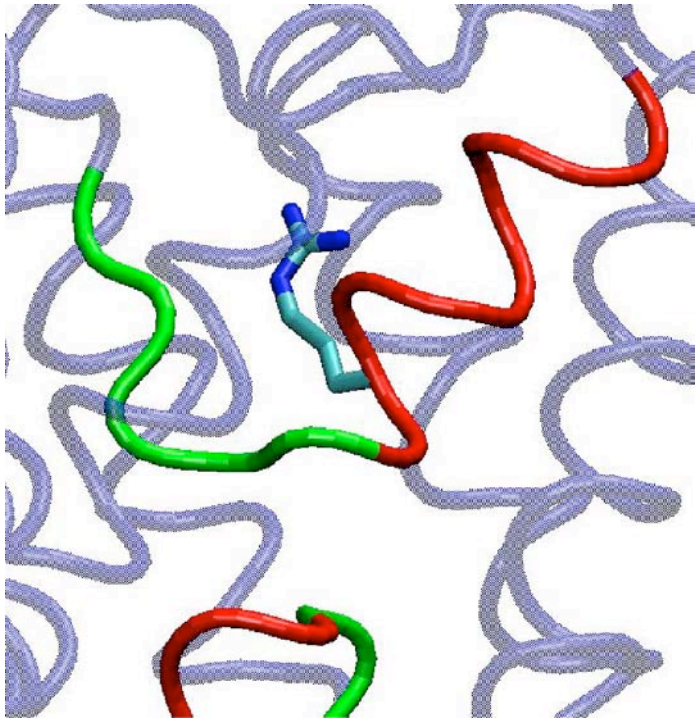
tube - transparent

Comparing Sequences of Aquaporins



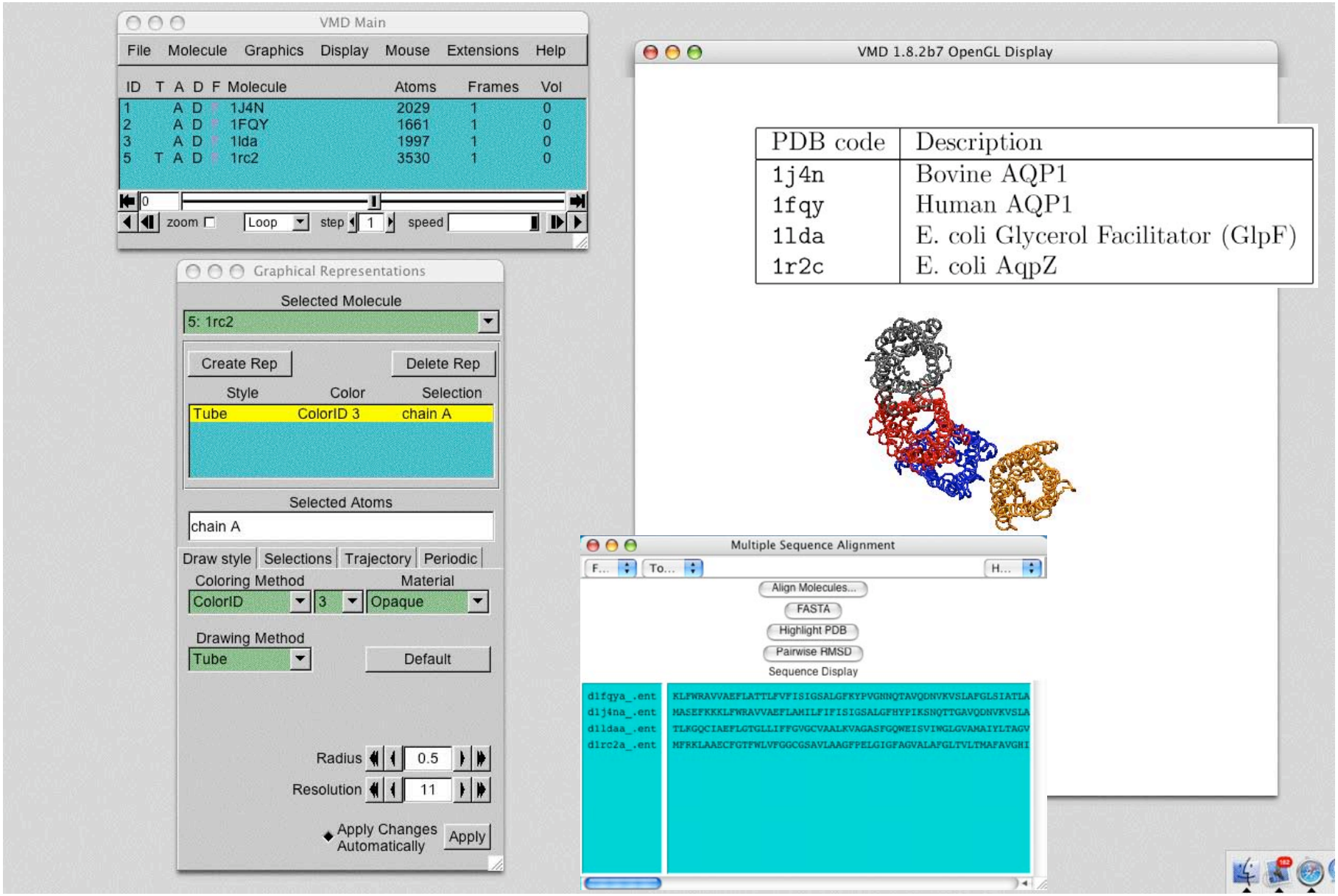
Highlighting Key Conserved Residues



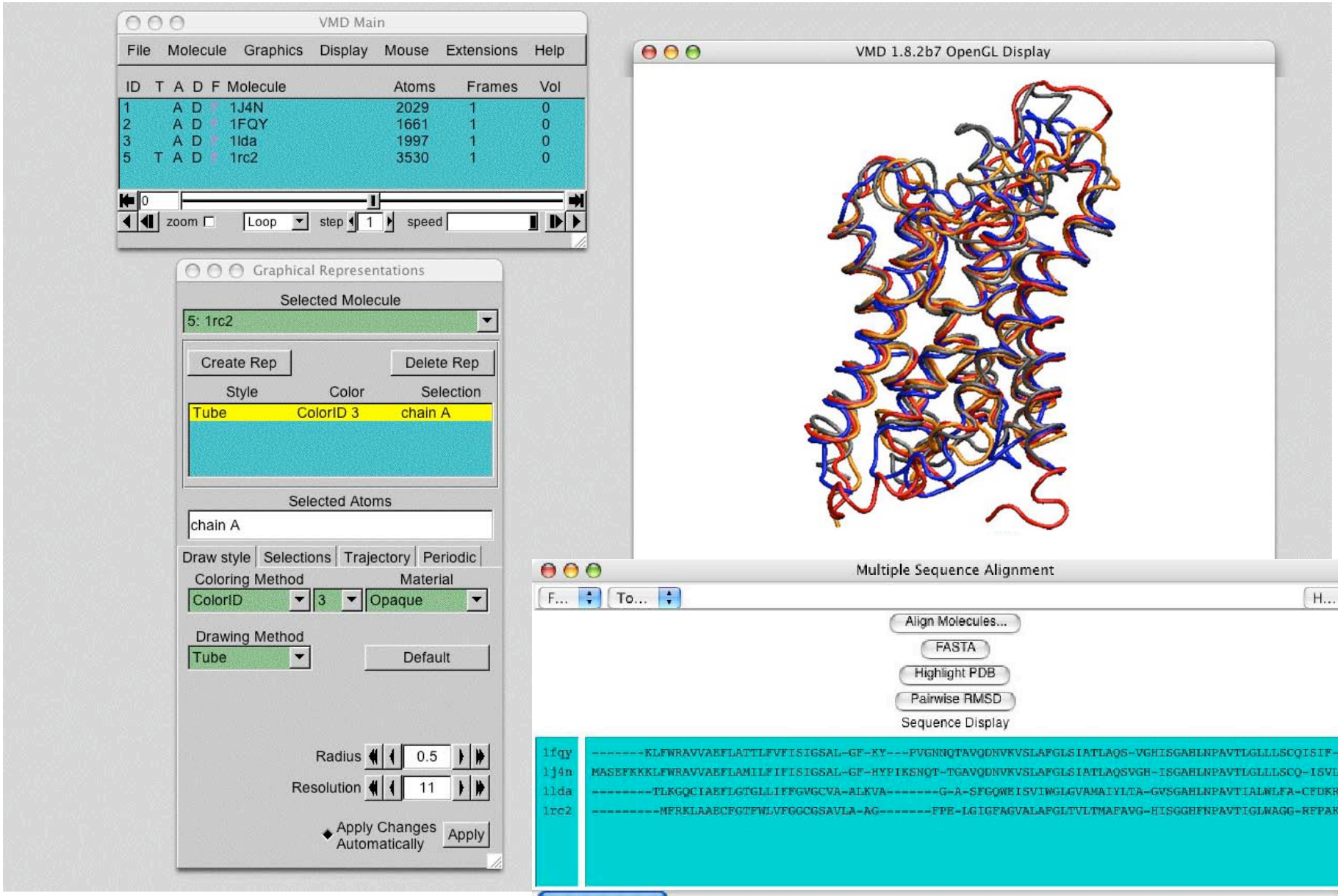


```
                 **:*  :  .  :.                        :   ..*:: ..          .:
AQP0_HUMAN       NPARSFAPAILTGNFT-------------NHWVYWVGPIIGGGLGSLLYDFLLFP-    225
AQP1_HUMAN       NPARSFGSAVITHNFS-------------NHWIFWVGPFIGGALAVLIYDFILAP-    233
AQP2_HUMAN       NPARSLAPAVVTGKFD-------------DHWVFWIGPLVGAILGSLLYNYVLFP-    225
AQP3_HUMAN       NPARDFGPRLFTALAGWGSAVFTTGQ--HWWWVPIVSPLLGSIAGVFVYQLMIGC-    267
AQP4_HUMAN       NPARSFGPAVIMGNWE-------------NHWIYWVGPIIGAVLAGGLYEYVFCPD    255
AQP5_HUMAN       NPARSFGPAVVMNRFSP------------AHWVFWVGPIVGAVLAAILYFYLLFP-    227
AQP6_HUMAN       NPARSFGPAIIIGKFT-------------VHWVFWVGPLMGALLASLIYNFVLFP-    237
AQP7_HUMAN       NPSRDLPPRIFTFIAGWGKQVFSNGE--NWWWVPVVAPLLGAYLGGIIYLVFIGS-    278
AQP8_HUMAN       NPARAFGPAVVANHWN-------------FHWIYWLGPLLAGLLVGLLIRCFIGD-    251
AQP9_HUMAN       NPARDLSPRLFTALAGWGFEVFRAGN--NFWWIPVVGPLVGAVIGGLIYVLVIEI-    268
GLPF_ECOLI       NPARDFGPKVFAWLAGWGNVAFTGGRDIPYFLVPLFGPIVGAIVGAFAYRKLIGR-    257
      ruler      ...250.......260.......270.......280.......290.......300
```

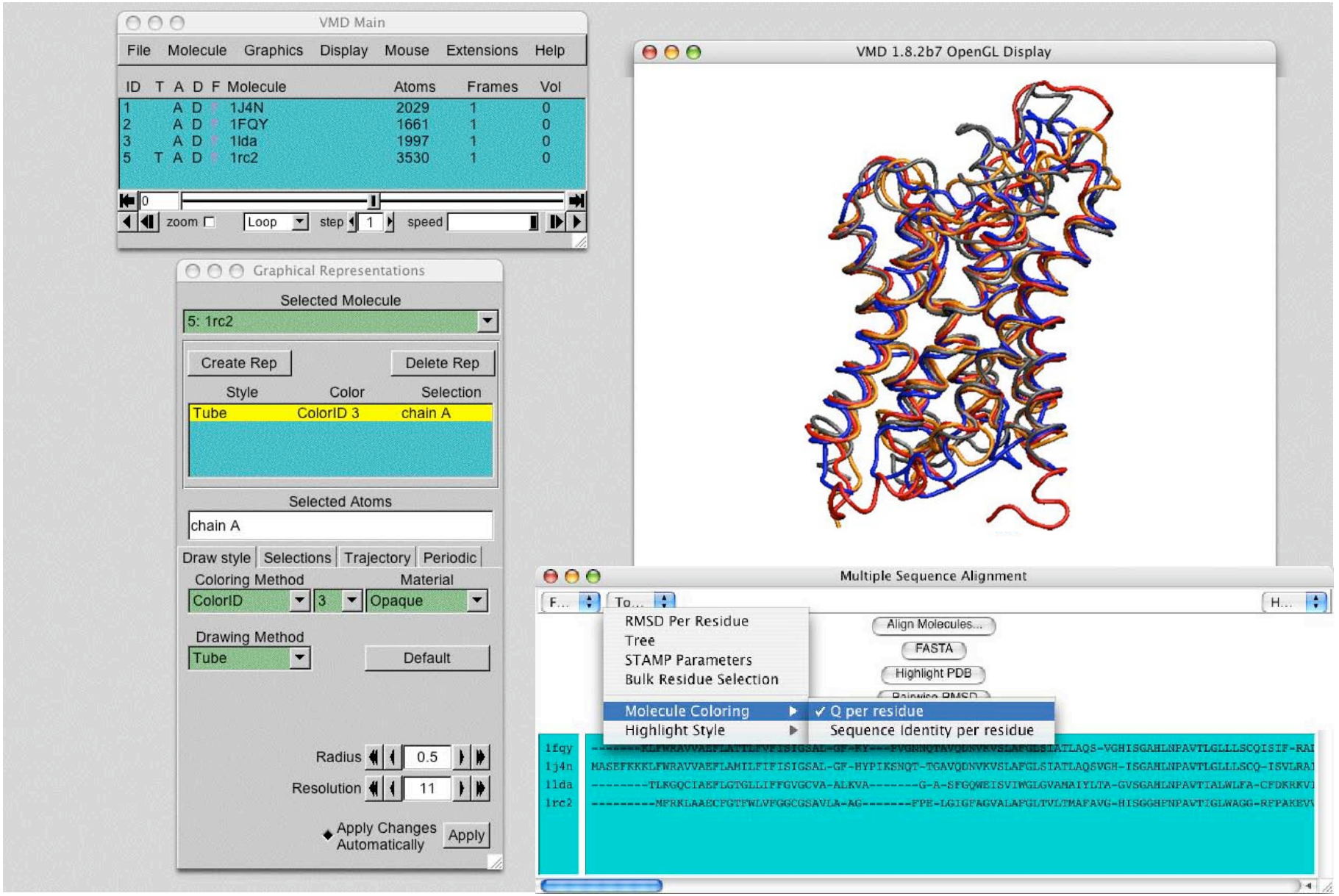
Load Aquaporins 1j4n, 1fqy, 1lda, 1rc2 into VMD



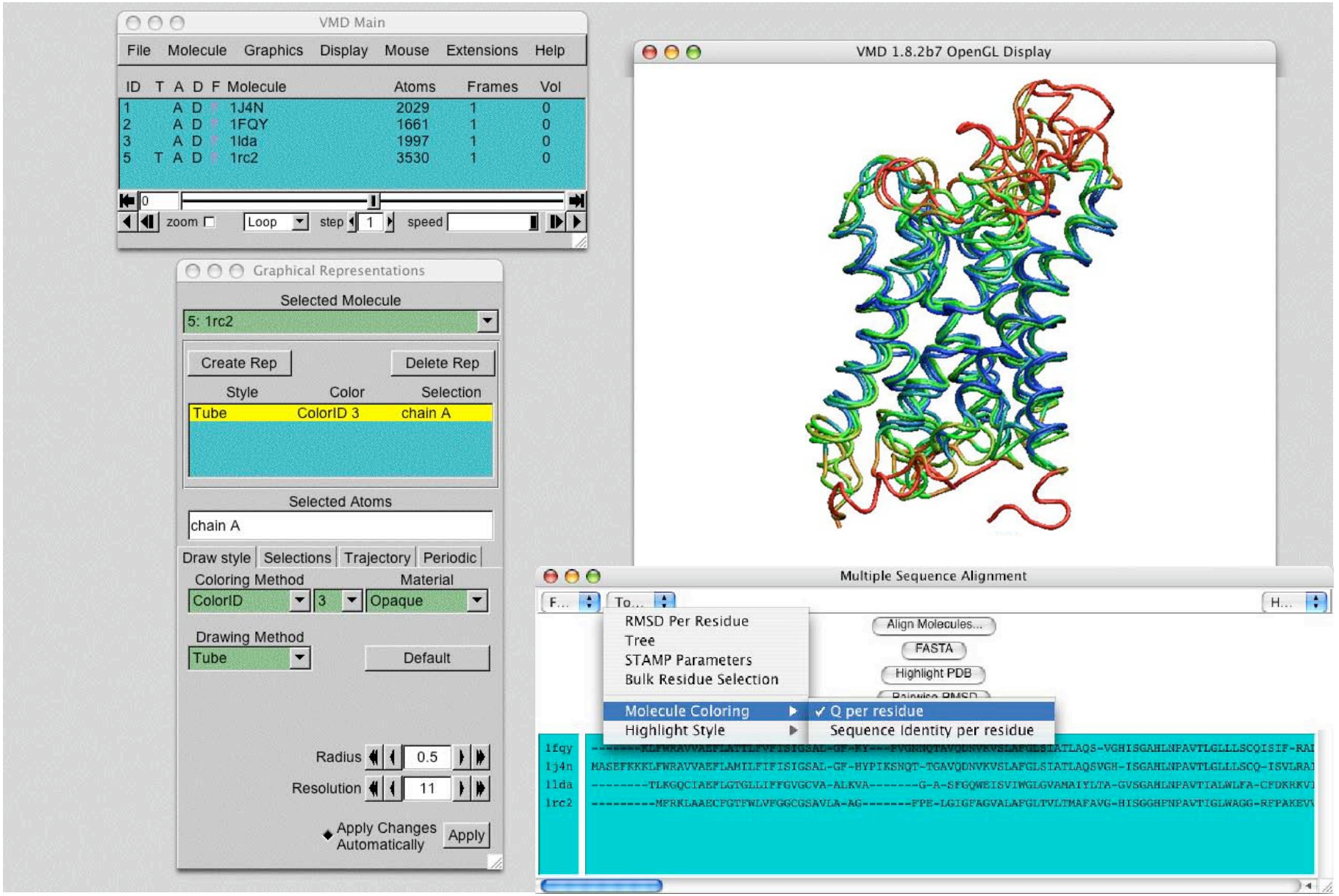
Aligning Structures and Sequences



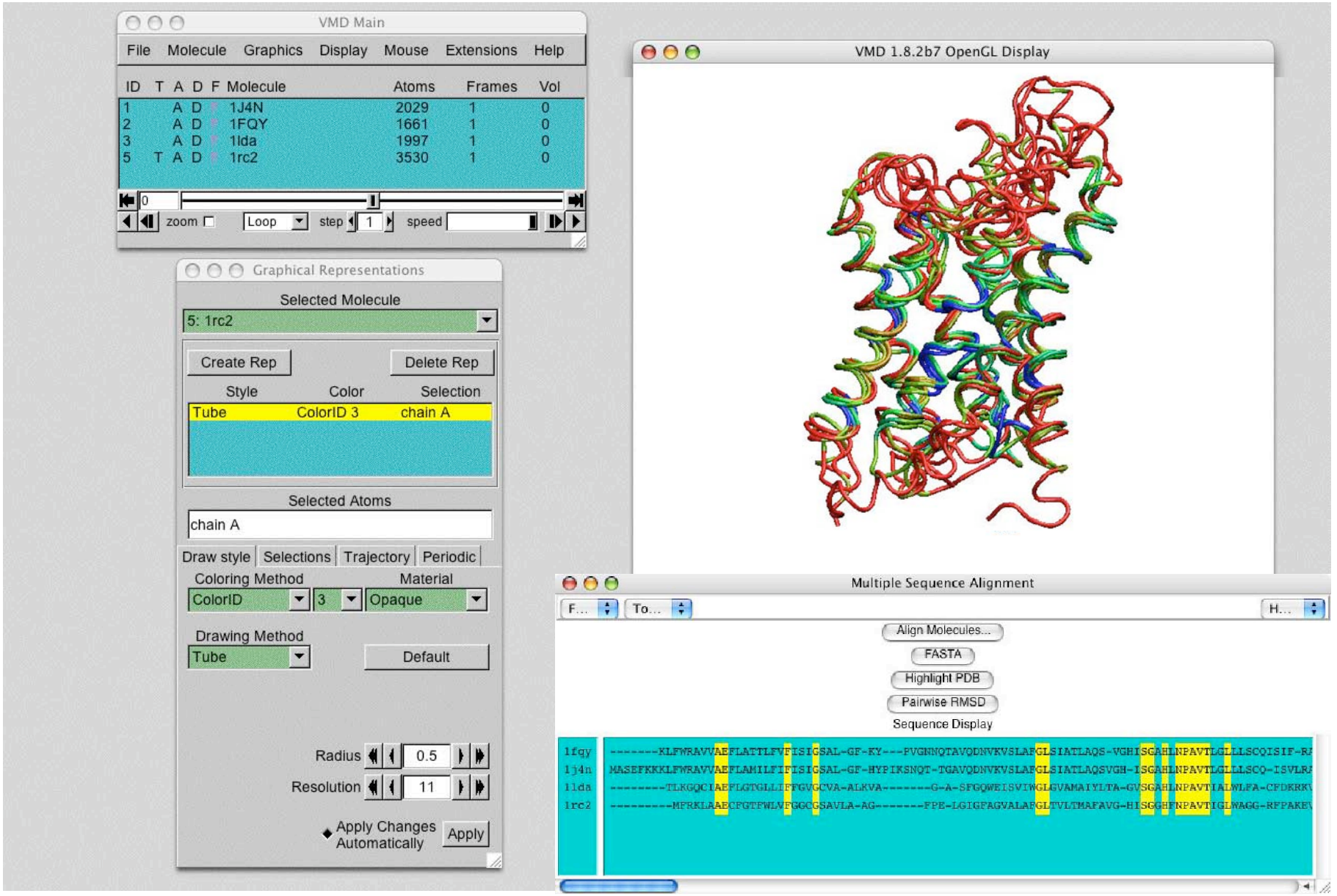
Comparing Structures by Similarity - Q Value



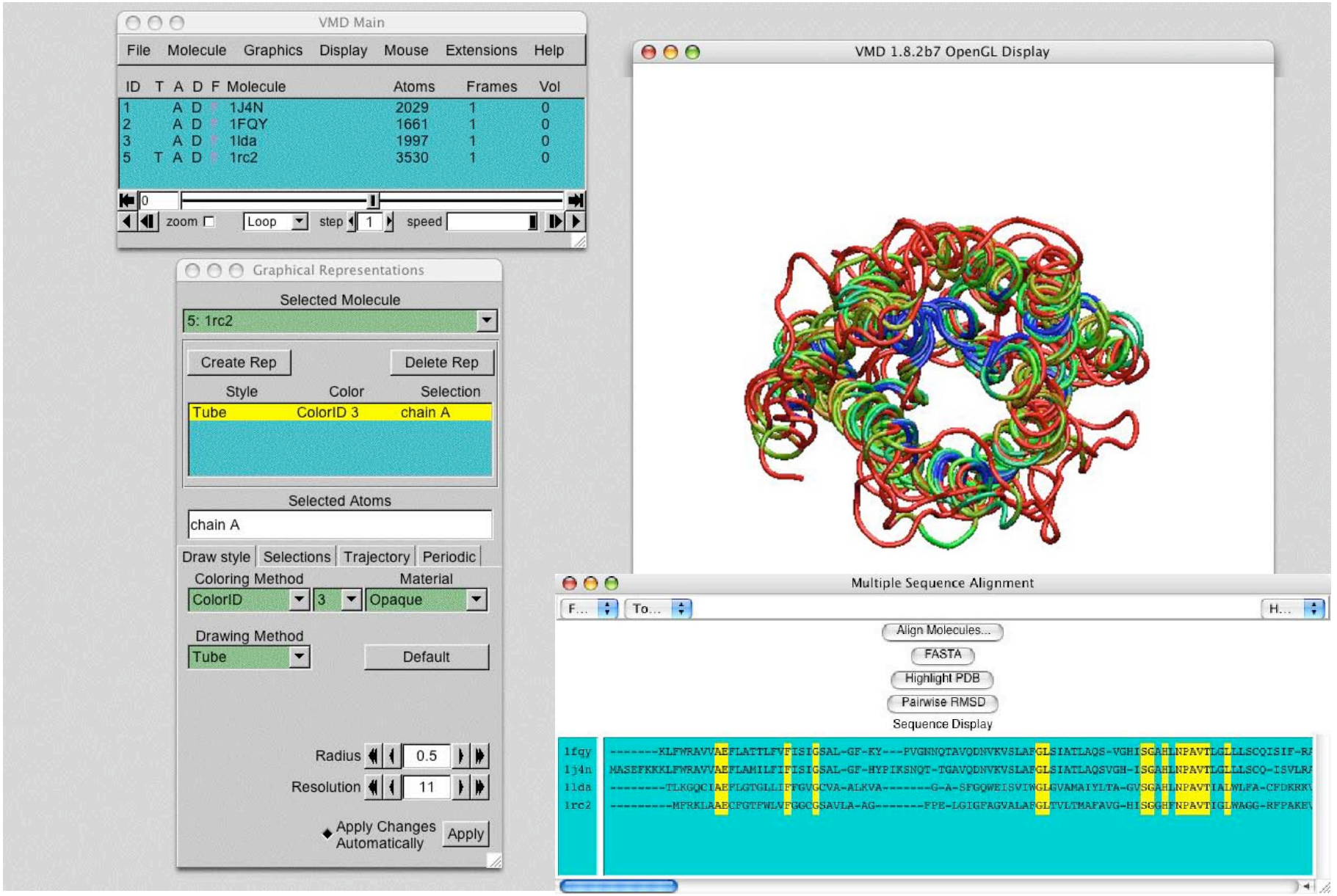
Comparing Structures by Similarity - Q Value



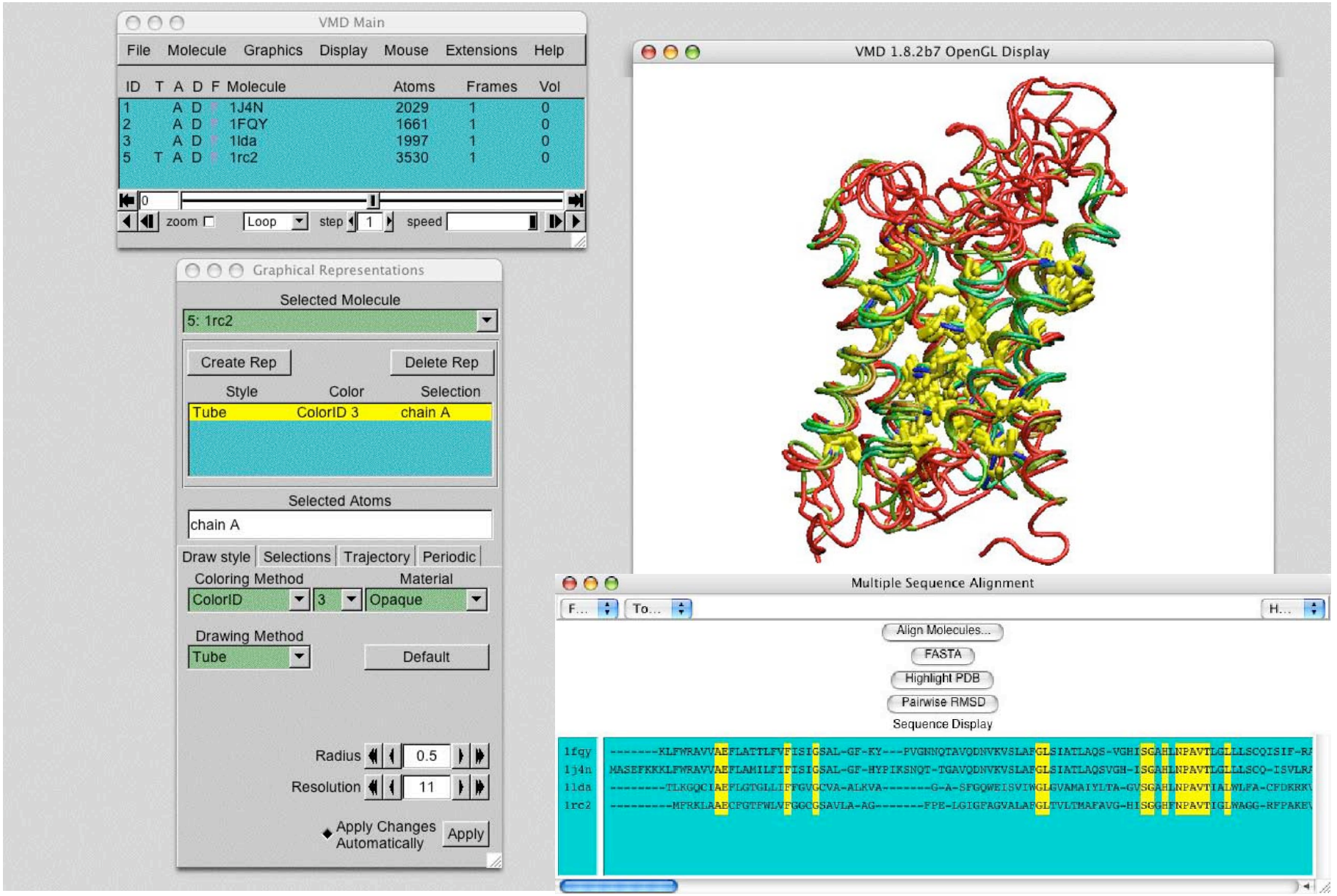
Exhibiting Sequence Identity - Side View



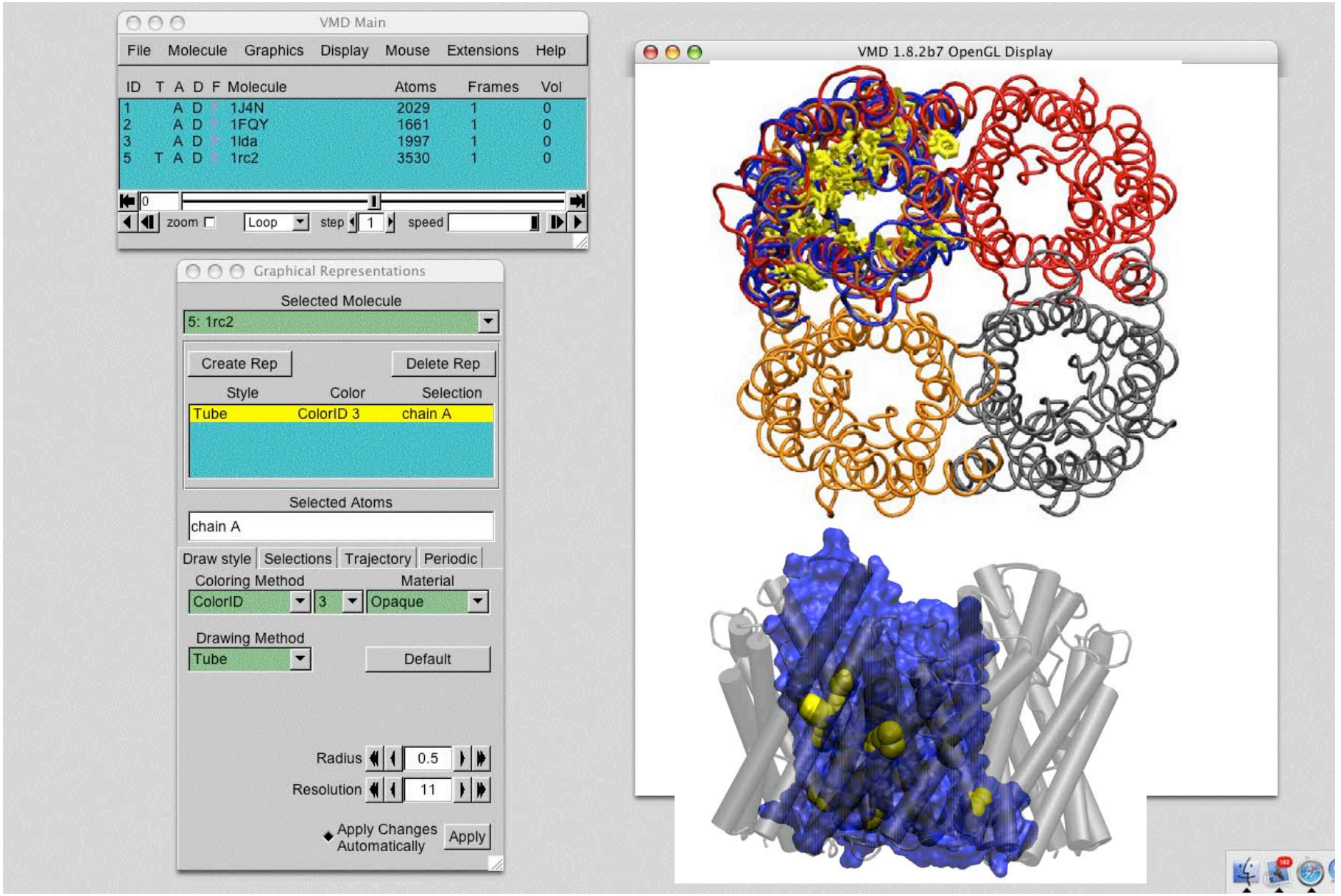
Exhibiting Sequence Identity - Top View



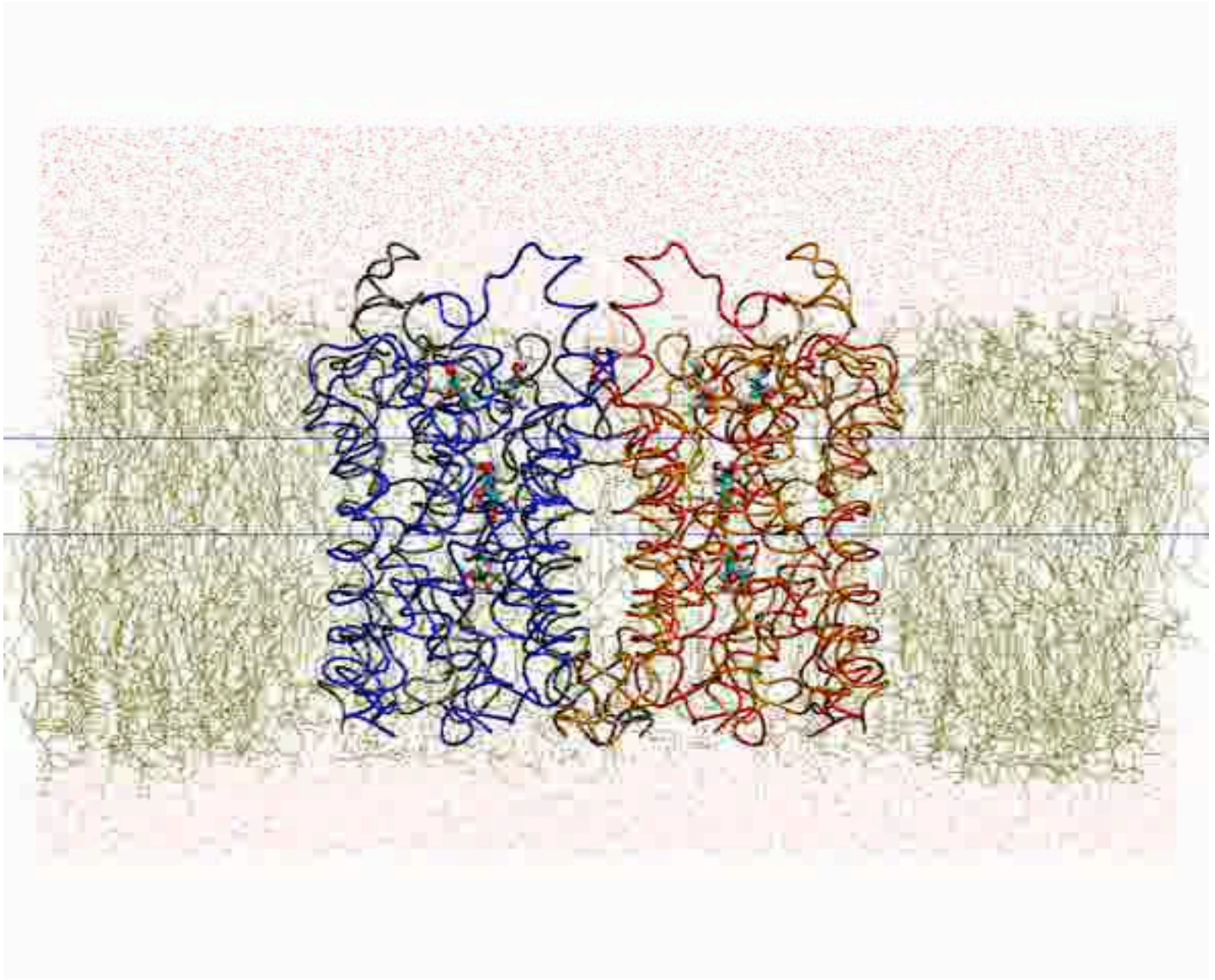
Showing Conserved Residues - Monomer



Showing Conserved Residues - Tetramer

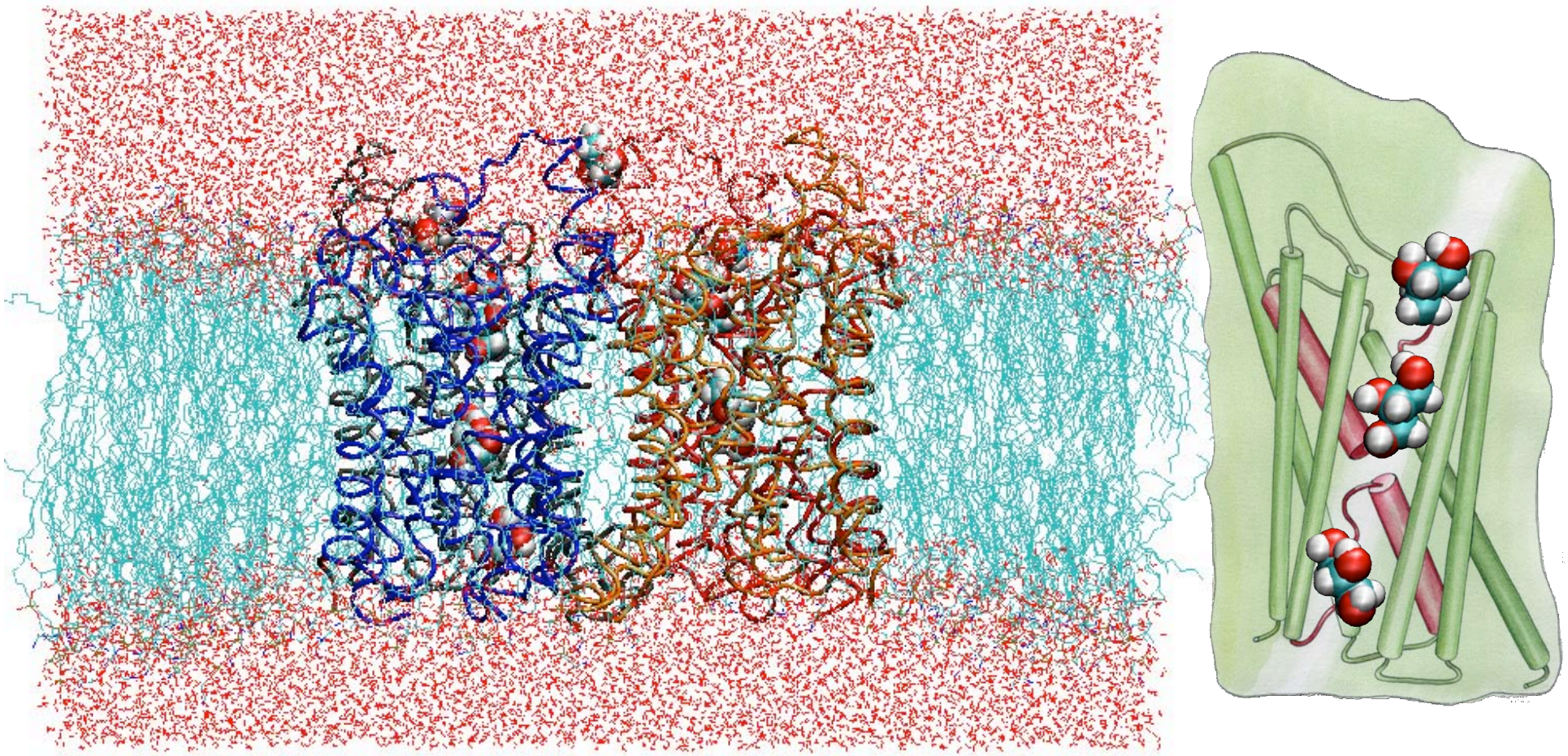


Dynamics of Protein, Lipid, Water System



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

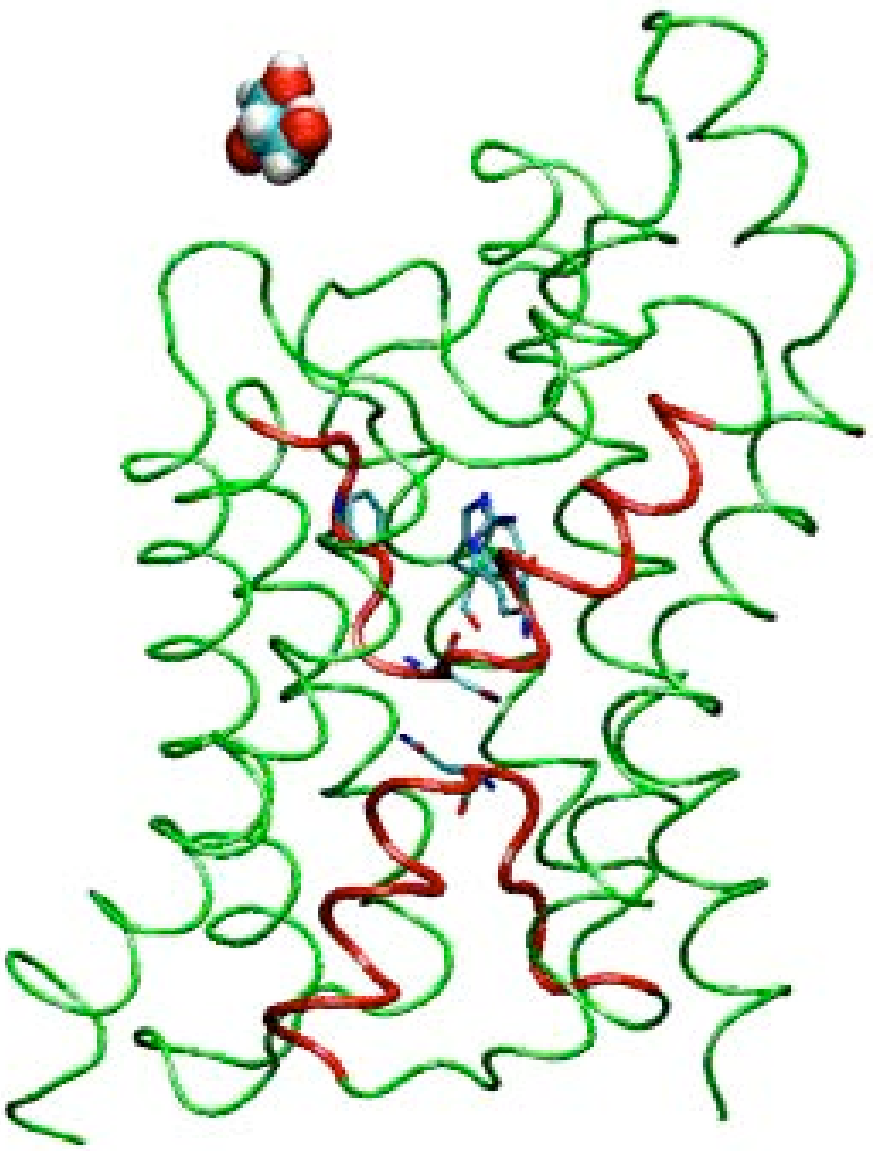
Equilibrated Structure after 1 ns



note the curved adjustment between lipids-protein

Morten Jensen, Emad Tajkhorshid

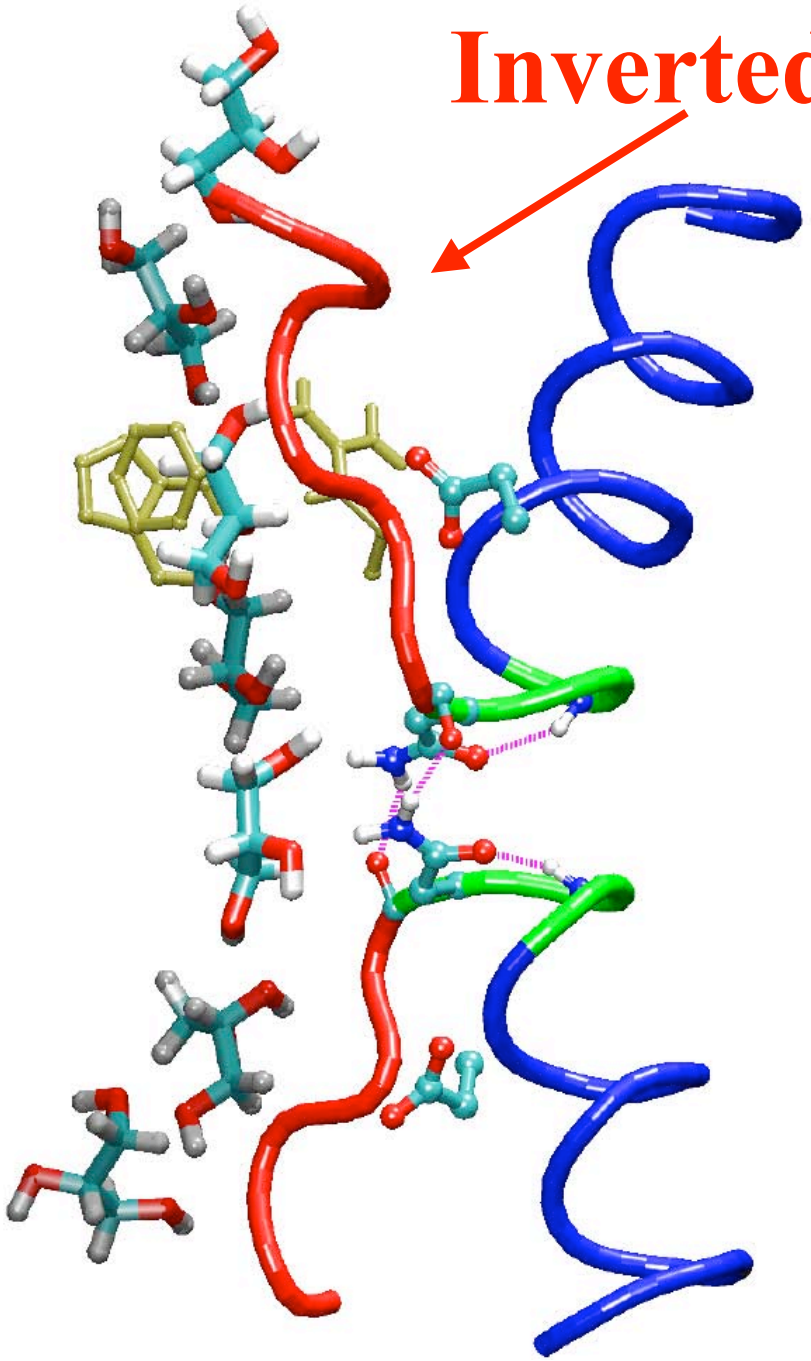
Glycerol Conduction



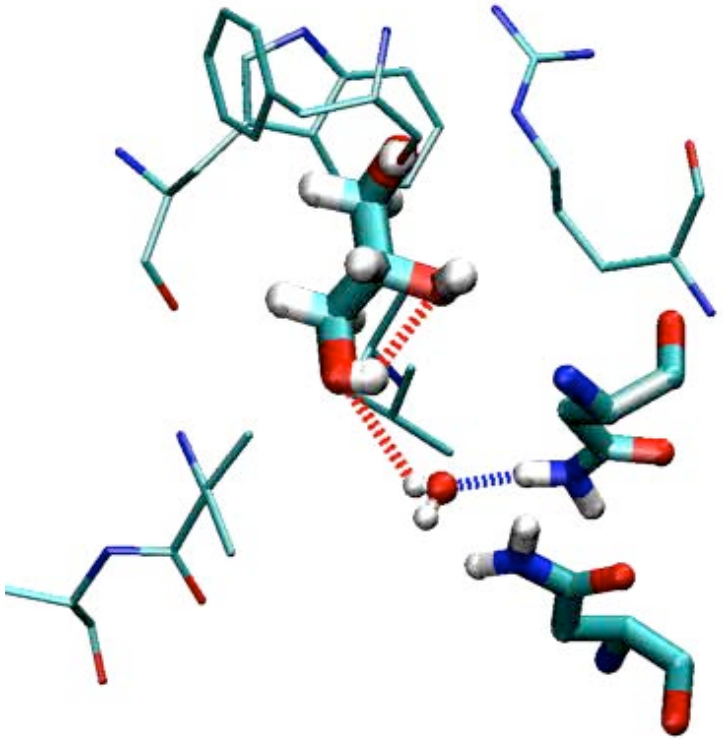
- Spontaneous glycerol conduction on ns time scale;
- Conduction occurs independently in each monomer;
- Exposed backbone carbonyl oxygen atoms dictates glycerol and water pathway; this explains the non-helical secondary structure in the aquaporin family;
- Glycerol resides at the positions of conserved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

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

Inverted helices guide glycerol



Glycerol – water competition for hydrogen bonds drives transport



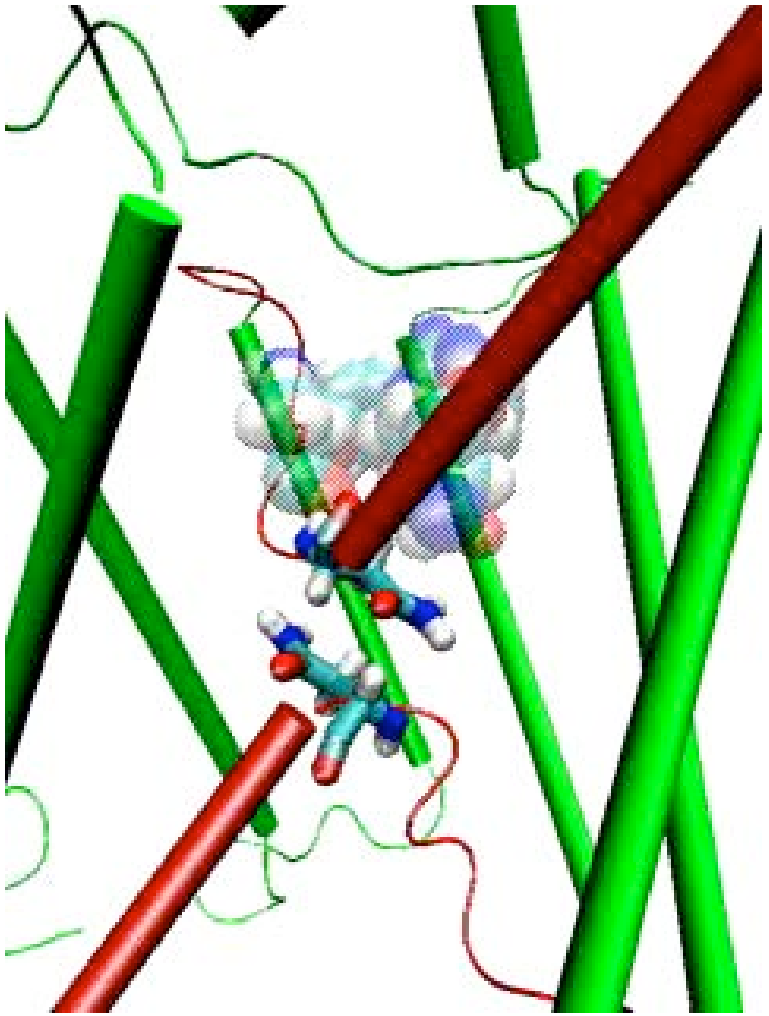
M. Jensen, E. Tajkhorshid, K. Schulten, *Structure* 9, 1083 (2001)

Interactive Molecular Dynamics

VMD ◄·····► **NAMD**

Molecular Graphics **Molecular Dynamics**

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



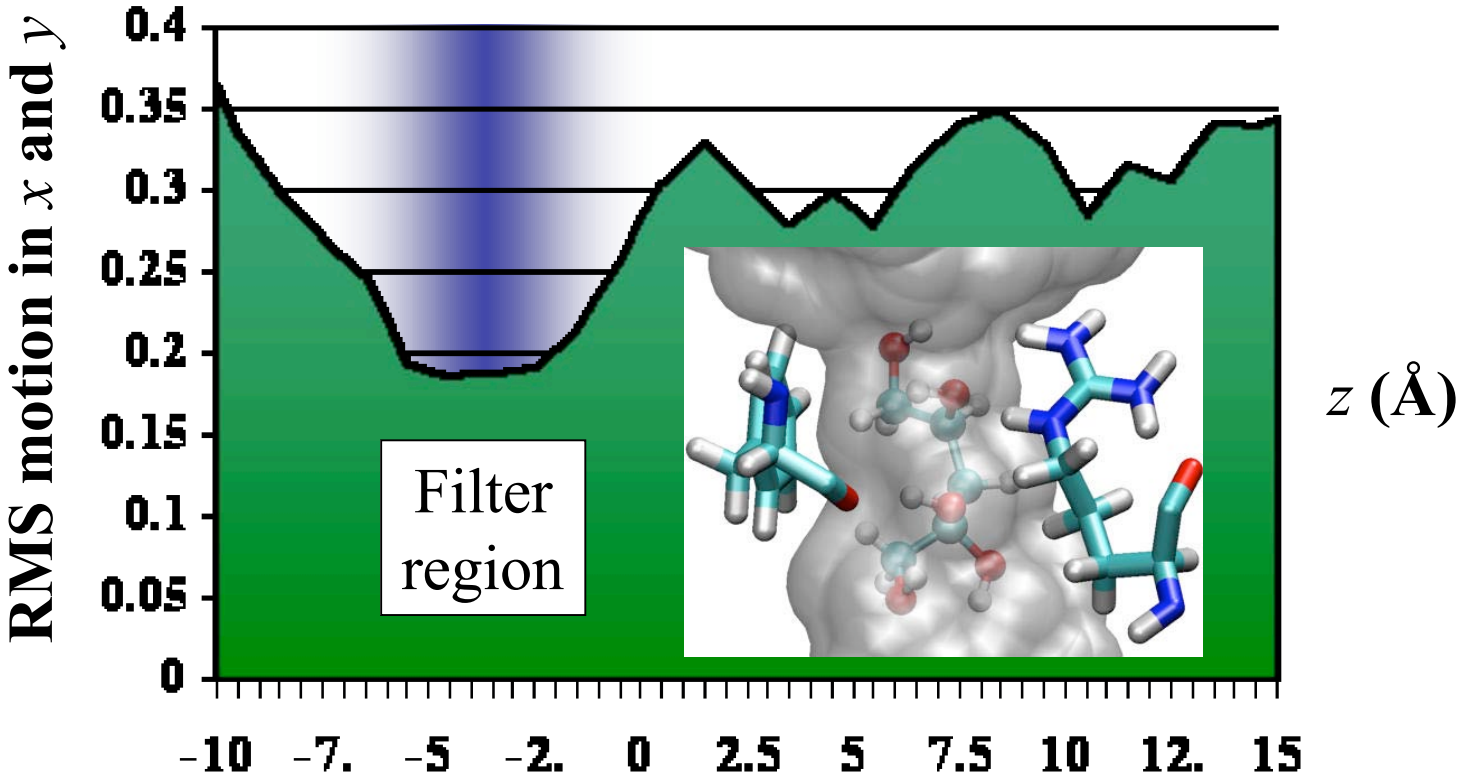
J. Stone, J. Gullingsrud, K. Schulten, and P. Grayson.
A System for Interactive Molecular Dynamics Simulation.
2001 ACM Symposium on Interactive 3D Graphics, pp.191-194, ACM SIGGRAPH
P. Grayson, E. Tajkhorshid, and K. Schulten.
Biophysical J, **83**: 36-48 (2003)

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

National Center for Research Resources

Confinement in Filter

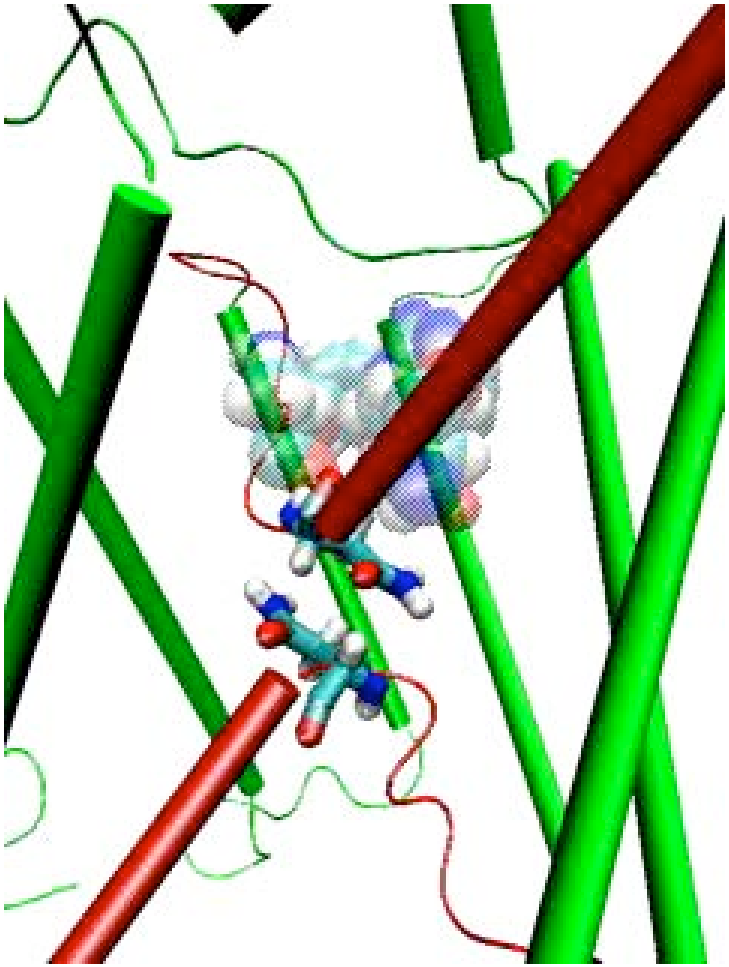
- **Selection occurs in most constrained region (induced fit)**
- **Selectivity probes shape, flexibility, hydrogen bonding.**

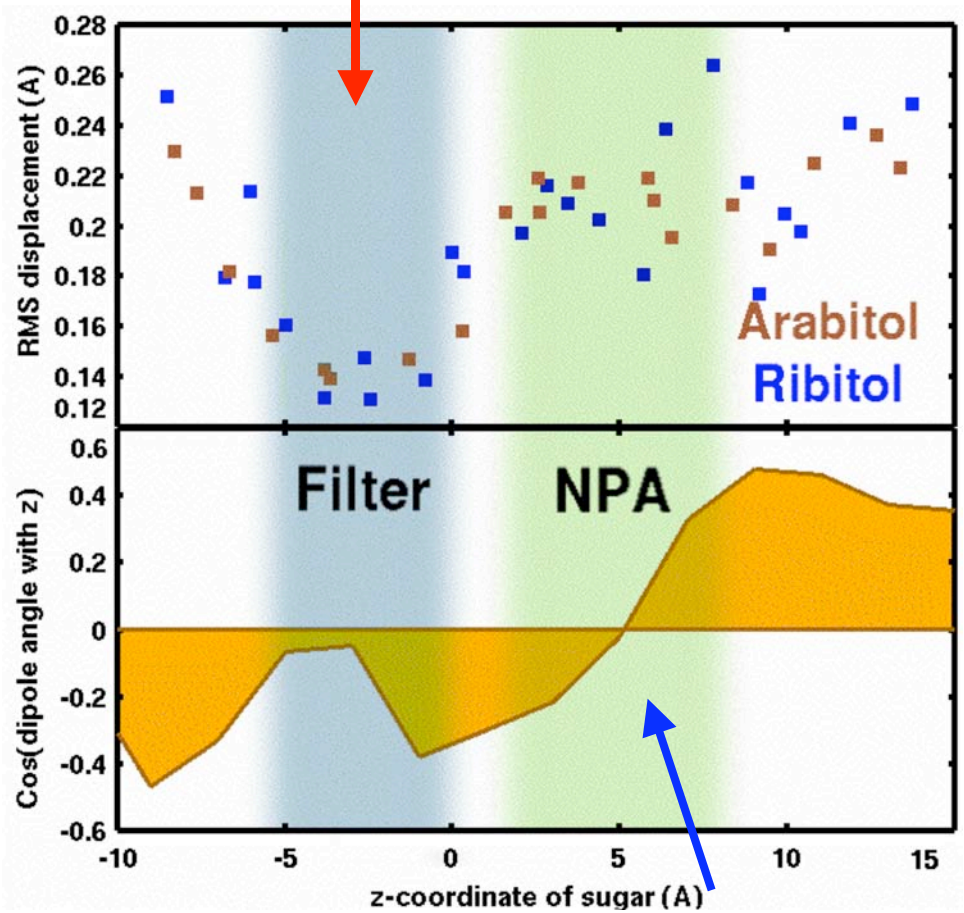


P. Grayson, E. Tajkhorshid, and K. Schulten. *Biophysical J.* **85**: 36-48 (2003)

Results of Interactive Simulations

Restricted motion filter for all sugars

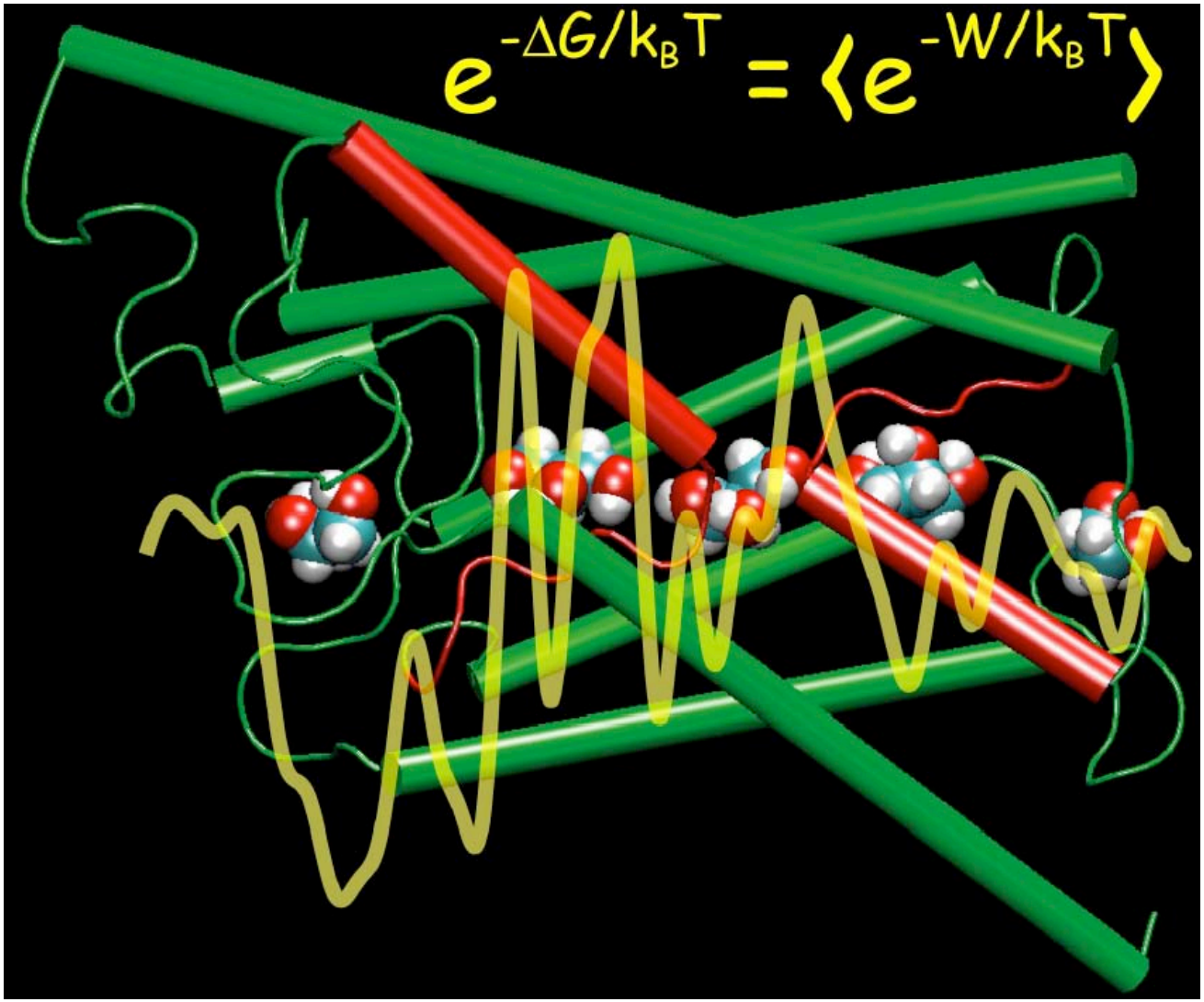




Dipole reversal at NPA

P. Grayson, E. Tajkhorshid, and K. Schulten. *Biophysical J.* **85**: 36-48 (2003)

Resulting Potential of Mean Force

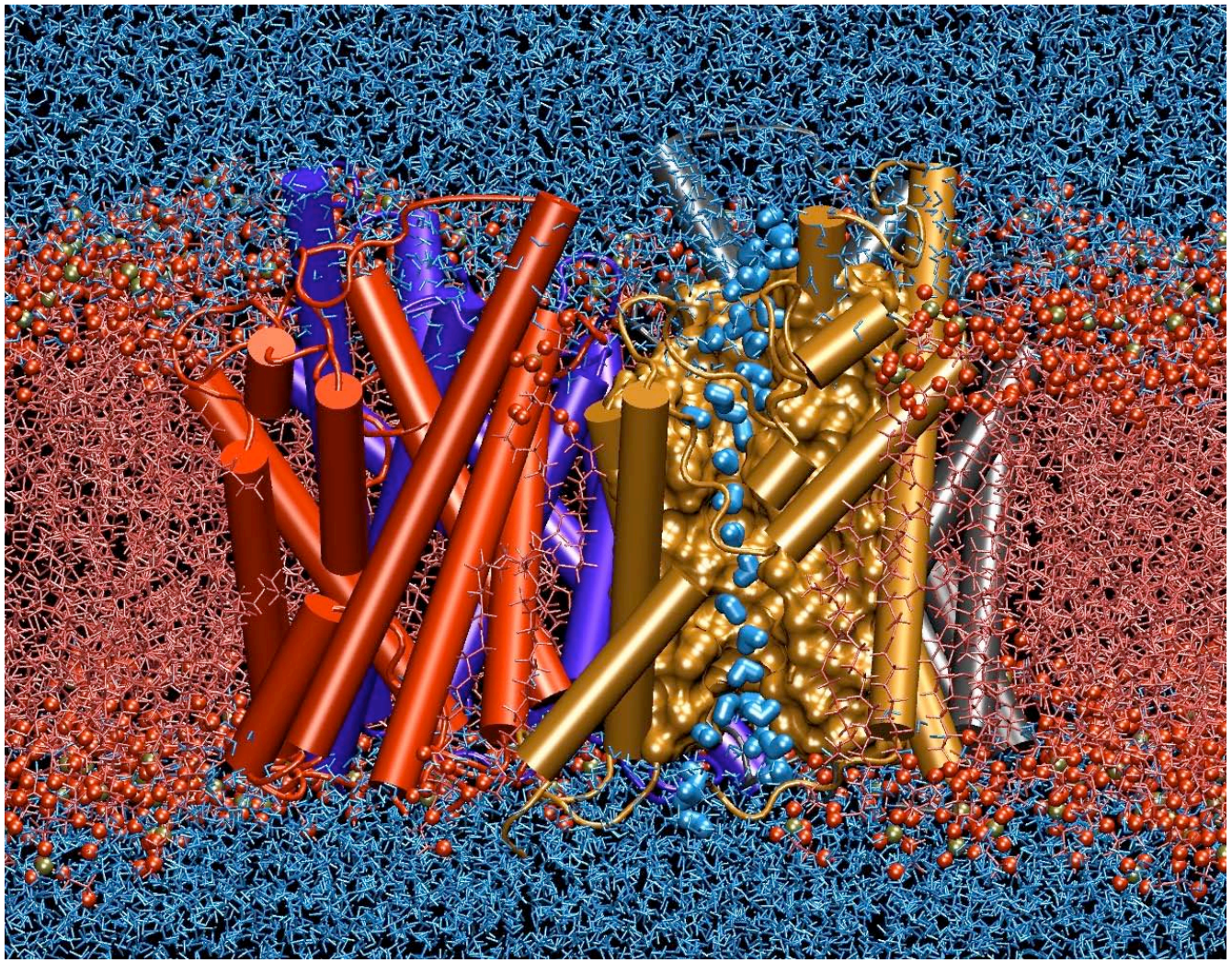


- The largest free energy barrier ≈ **7.3 kcal/mol**

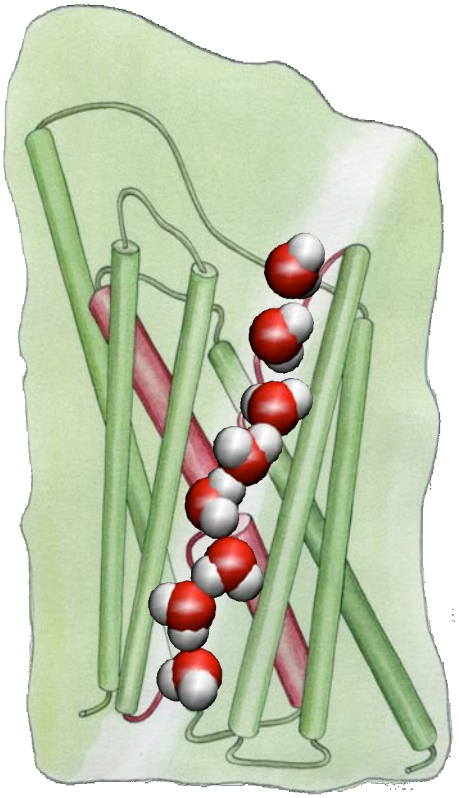
  cf. Arhenius activation energy measured:
  9.6±1.5 kcal/mol, Borgnia and Agre (2001)

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

Simulated System 3: GlpF With Only Water

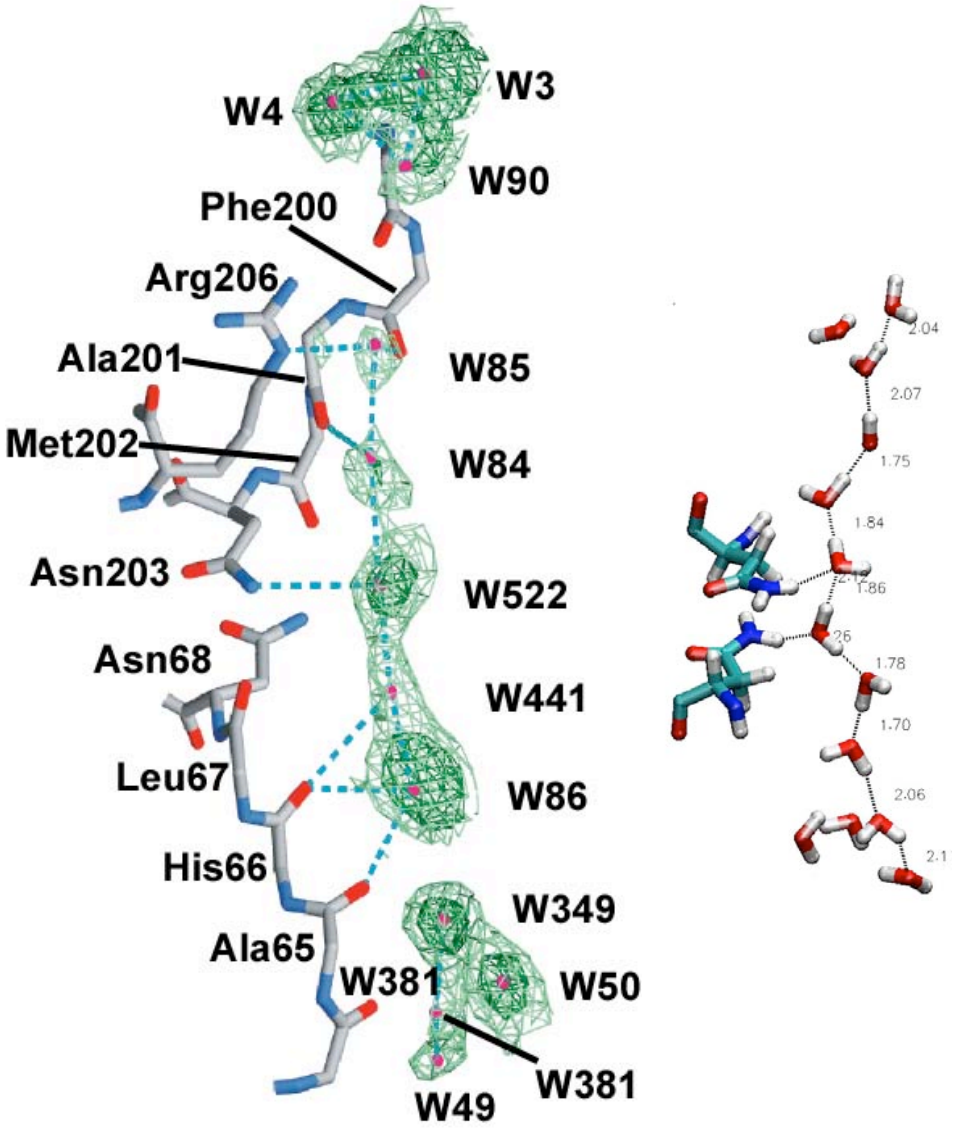


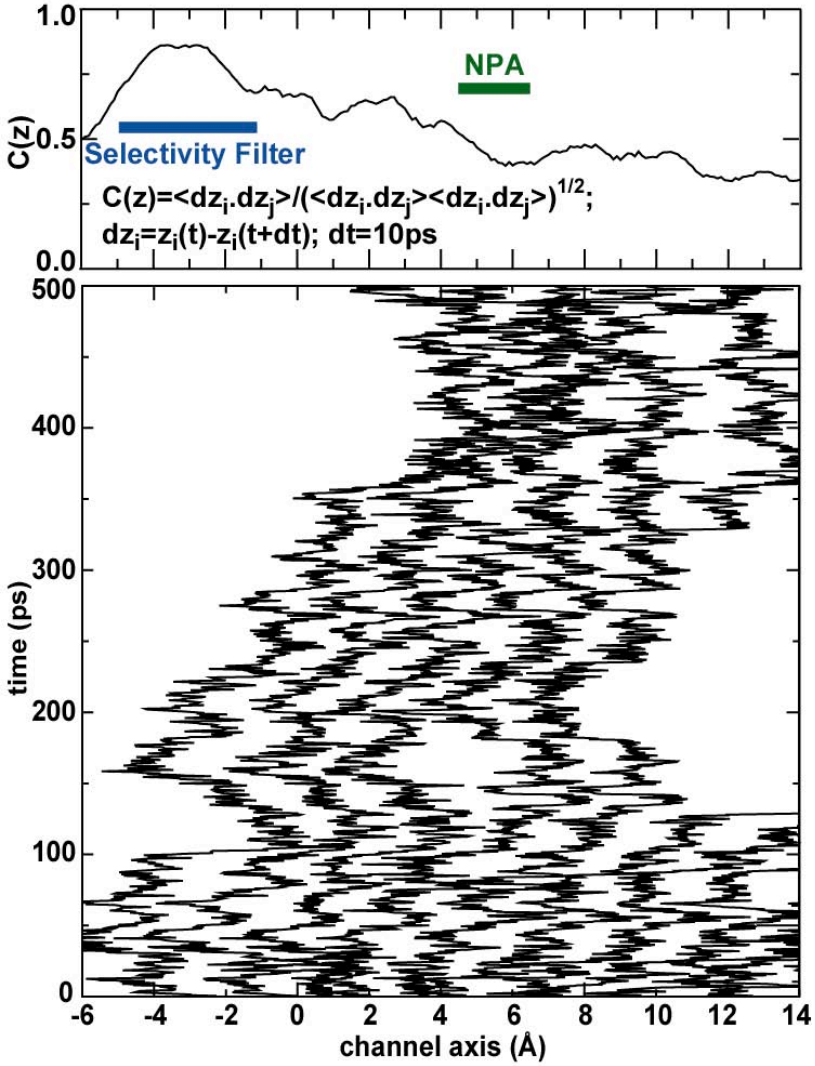
5ns + 2ns simulation



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

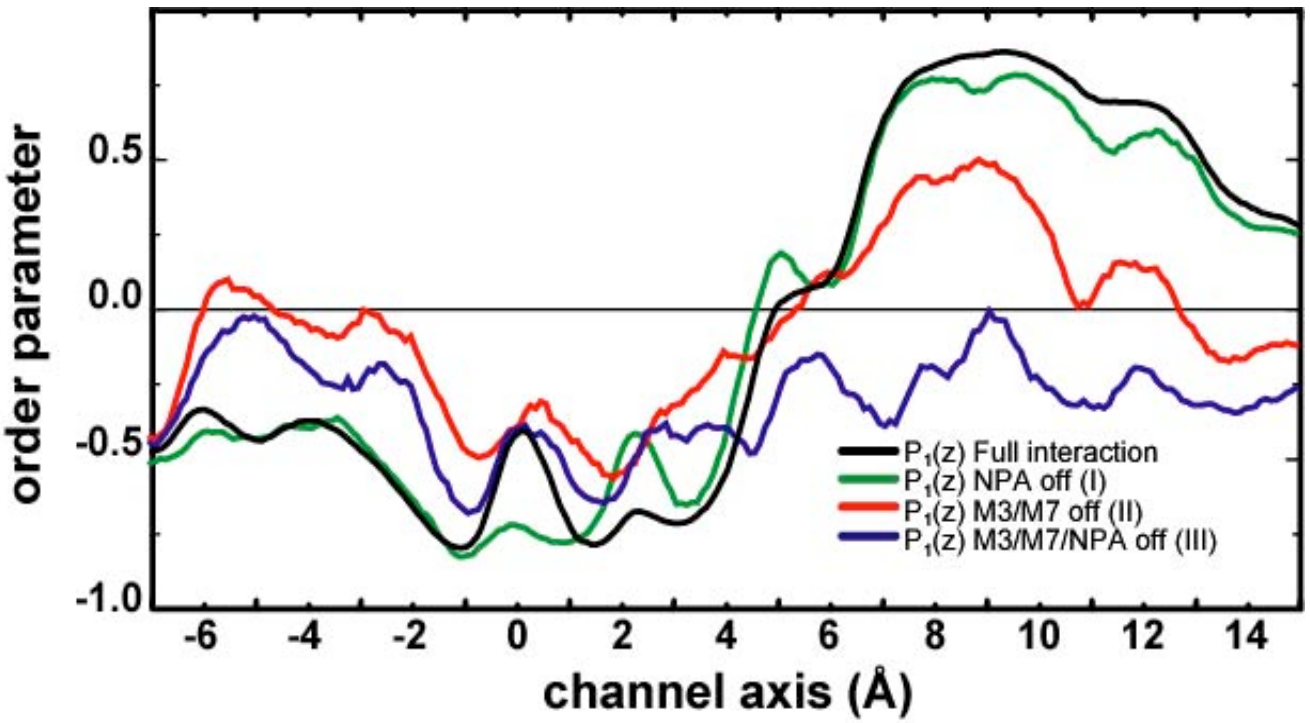
Water Positions Determined by Simulations and Experiments

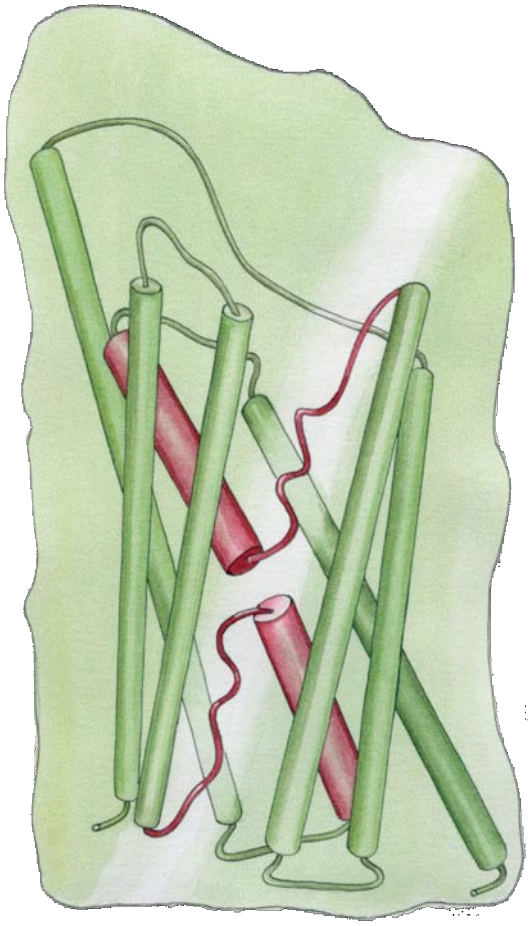




E. Tajkhorshid, P. Nollert, M. Jensen, L. J. W. Miercke, J. O'Connell, R. M. Stroud, and K. Schulten, ***Science*** 296, 525-530 (2002)

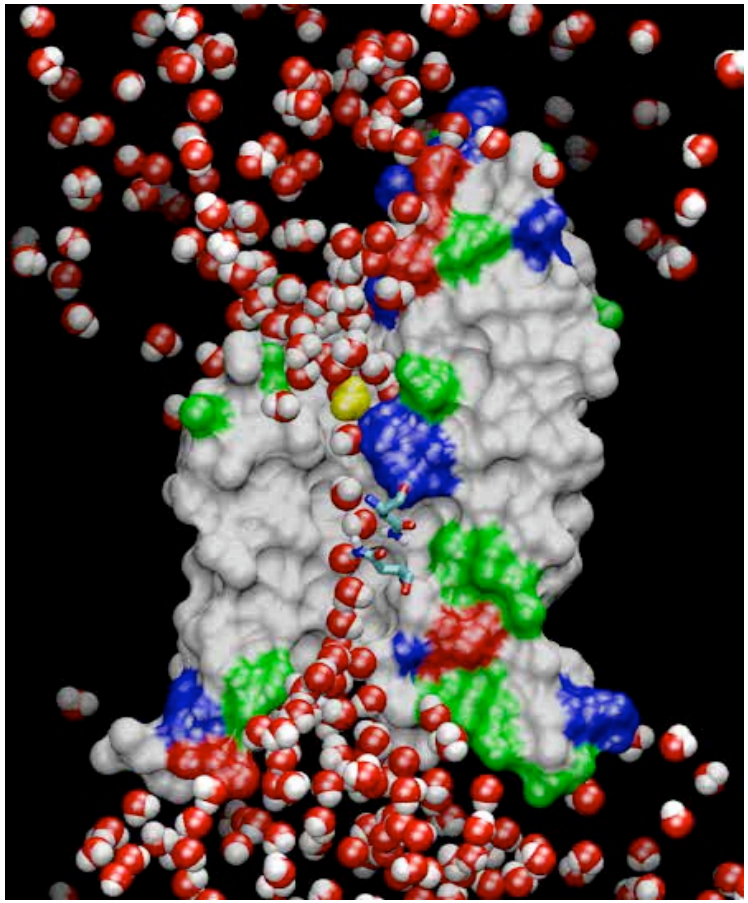
Electrostatic Stabilization of Water Bipolar Arrangement

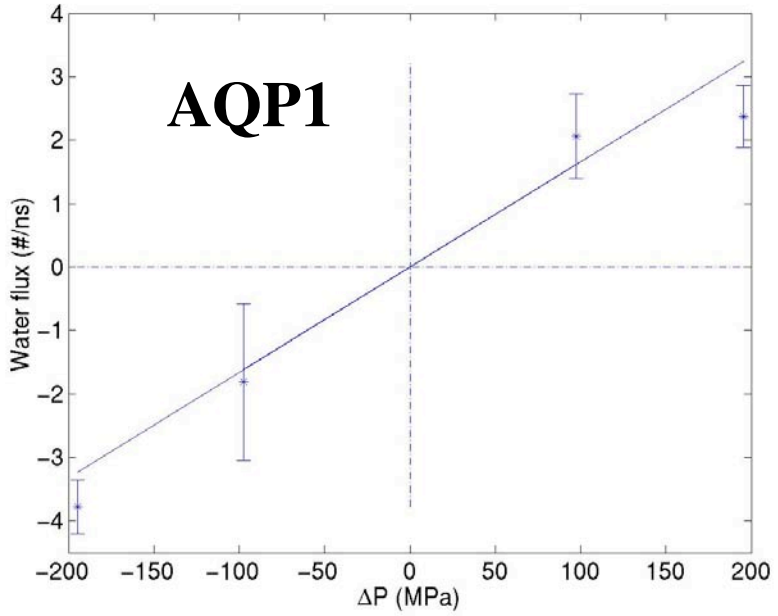




E. Tajkhorshid, P. Nollert, M. Jensen, L. J. W. Miercke, J. O'Connell, R. M. Stroud, and K. Schulten, ***Science*** 296, 525-530 (2002)

Water flux *vs.* pressure difference





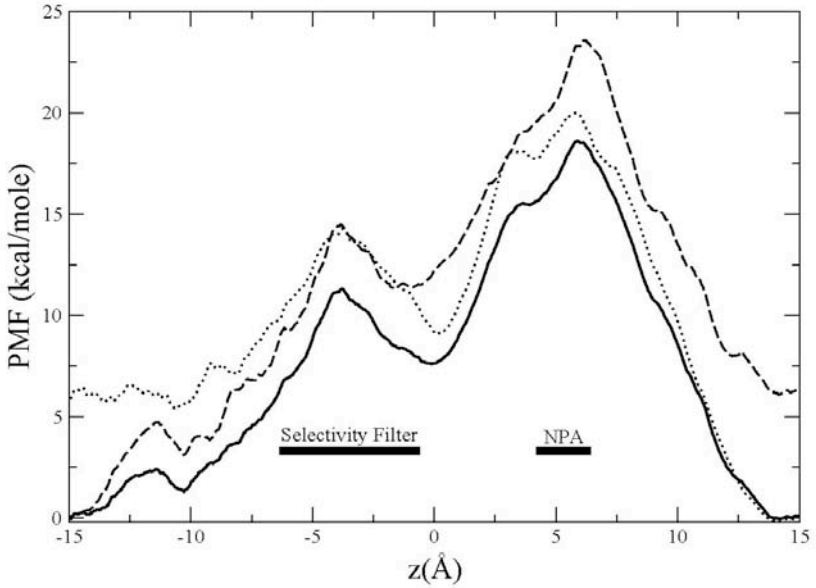
- Calculated p_f: $(7.1\pm0.9)\times10^{-14}$ cm^3/s
- Experimental p_f values: $5\sim11\times10^{-14}$ cm^3/s

F. Zhu, E. Tajkhorshid, and K. Schulten, ***Biophysical J.*** 86, 50-57 (2004)

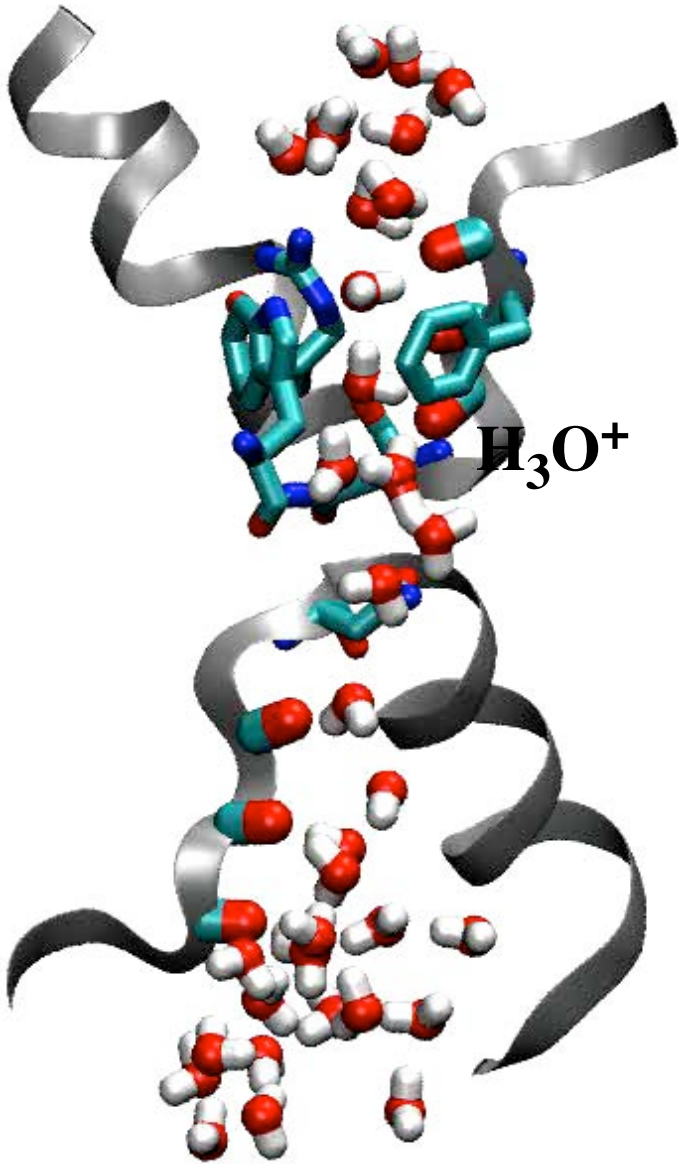
Proton Exclusion in Aquaporin Channel

Initial condition: H_3O^+ in center

Energy barrier preventing proton conduction



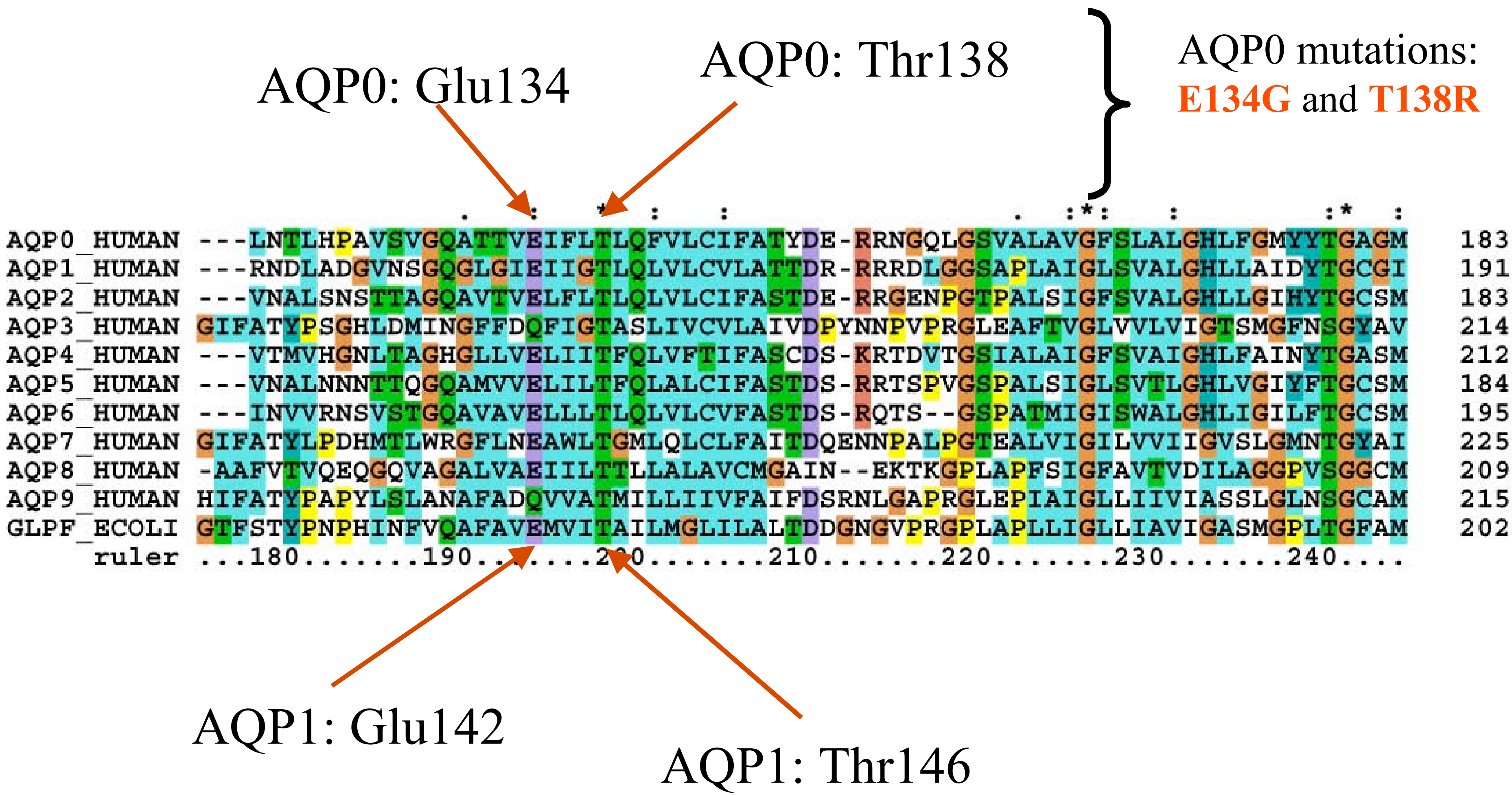
B. Ilan, E. Tajkhorshid, K. Schulten, G. Voth, *PROTEINS* **55**: 223-228 (2004)



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

Genetically Inherited Cataracts

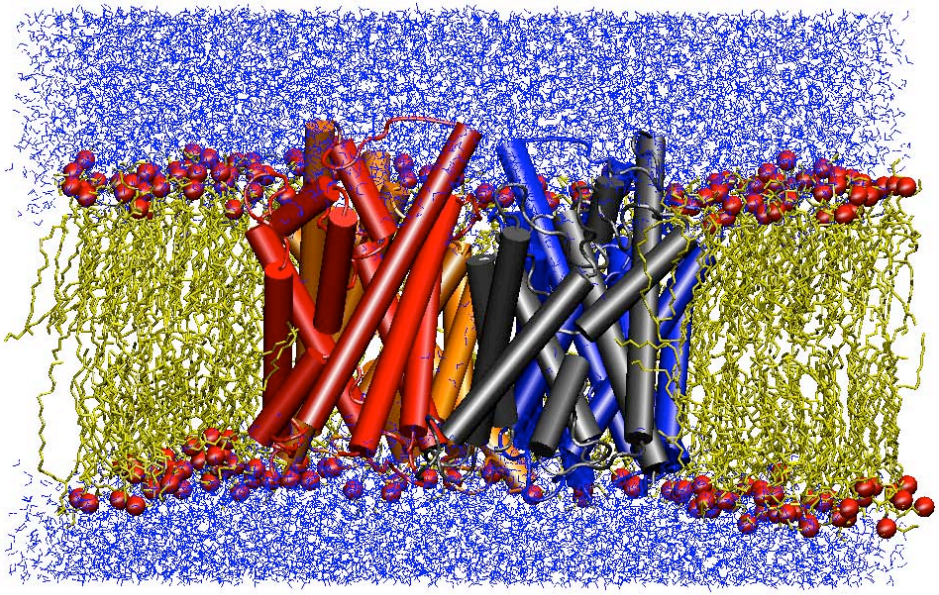
Impaired protein trafficking is suggested to be the main effect of these mutants, however, an impaired channel activity can also be involved.

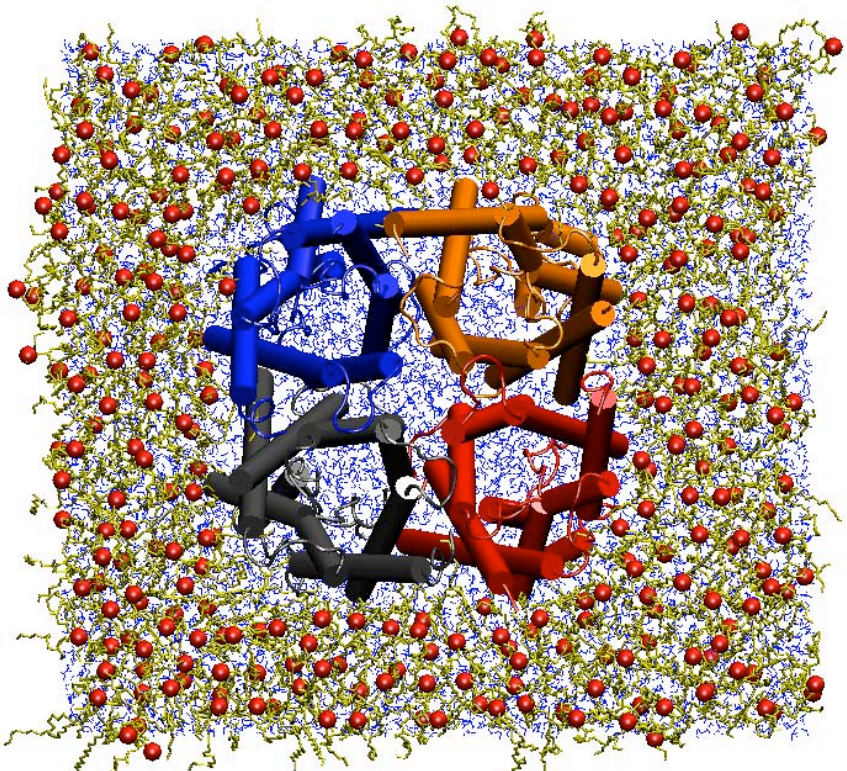


E. Tajkhorshid et al (unpublished)

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

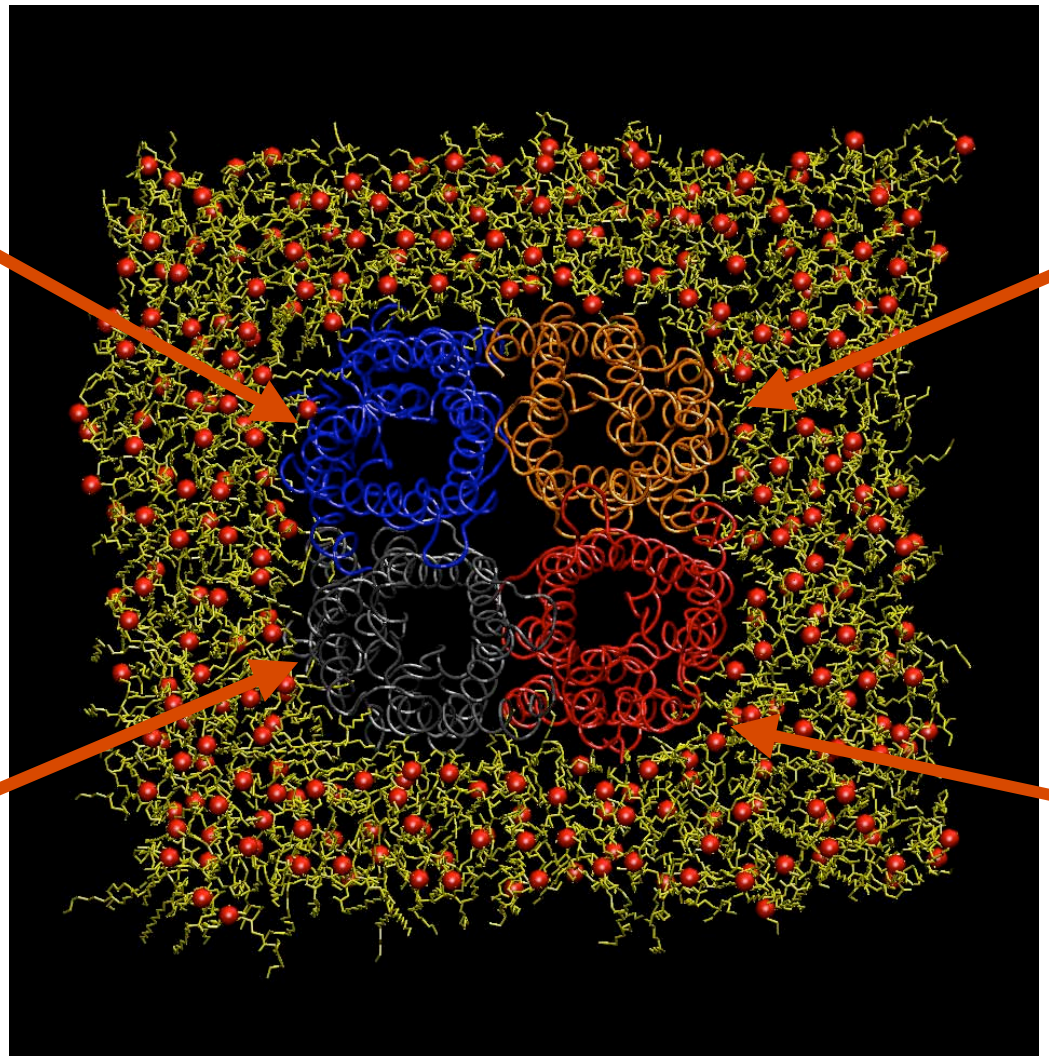
5 ns run of wild-type protein

15 ns simulation after mutation

10 days /ns – 32-proc Linux cluster

3.5 days/ns - 128 O2000 CPUs

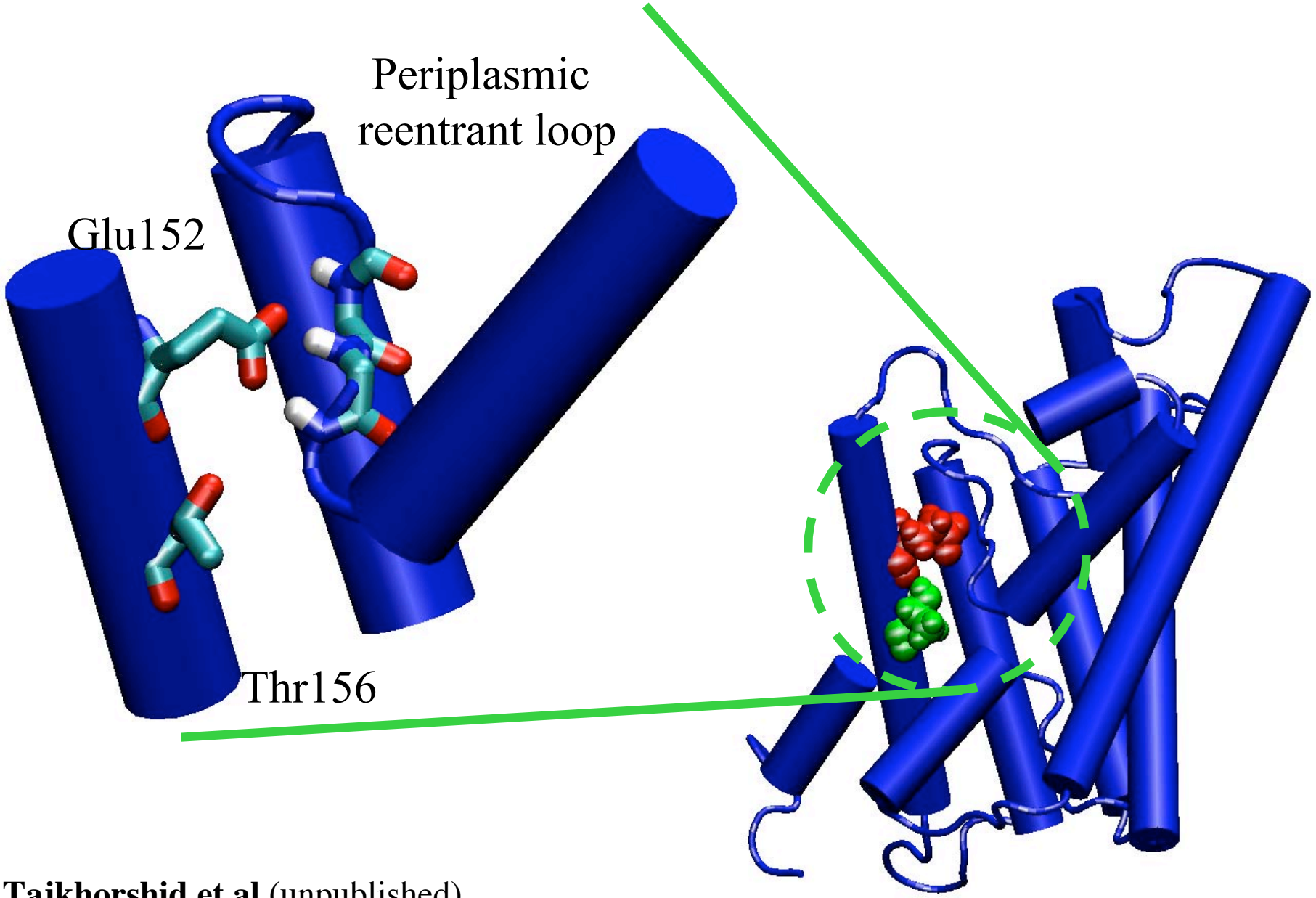
Point Mutations in the Tetramer



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

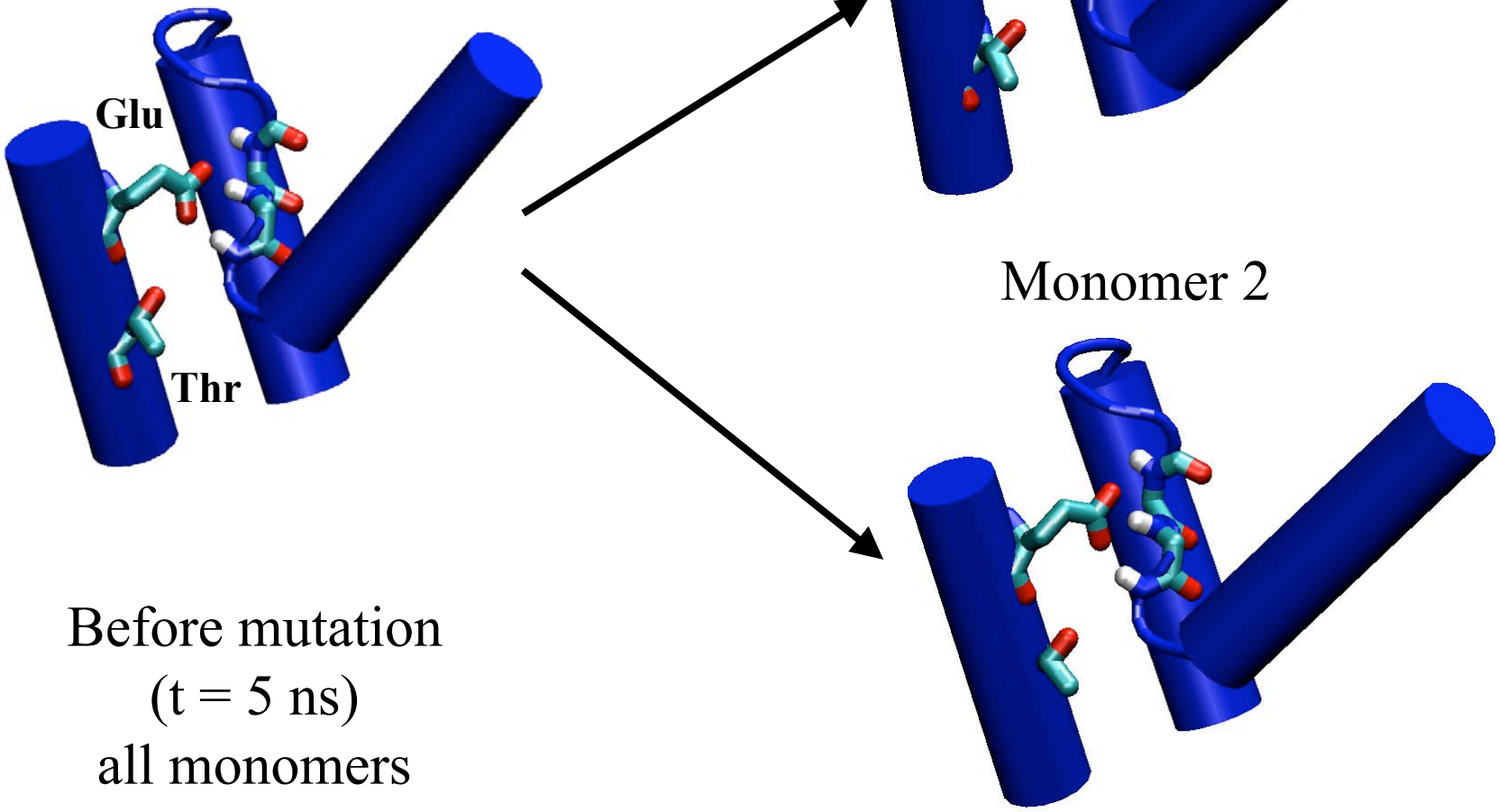
E. Tajkhorshid et al (unpublished)

Behaviour of Wild Type Aquaporin



E. Tajkhorshid et al (unpublished)

Wildtype Conformational Changes

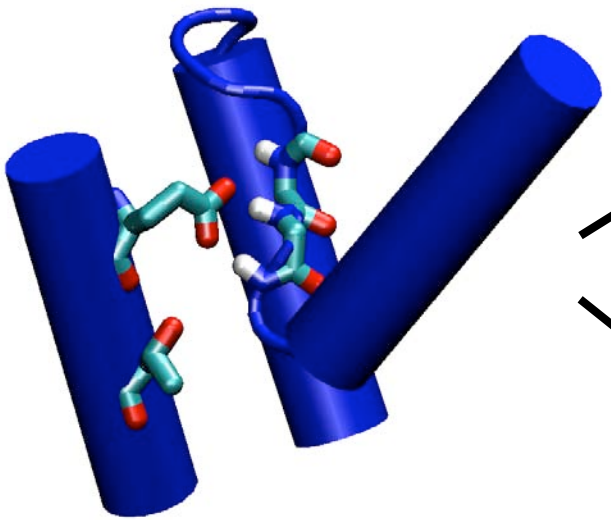


Before mutation
(t = 5 ns)
all monomers

t = 20 ns

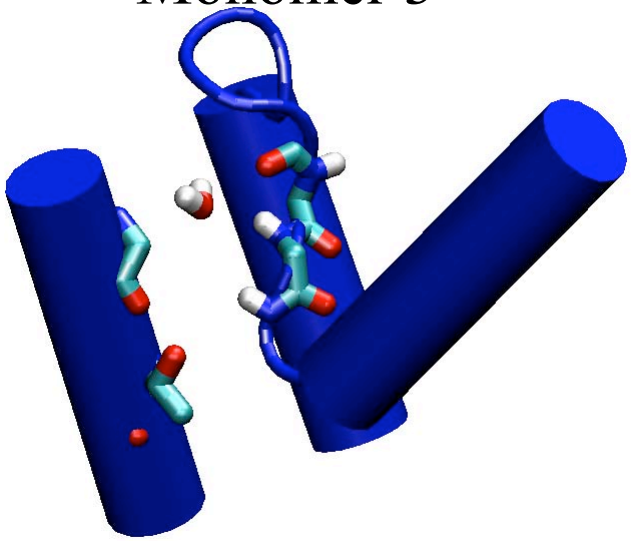
E. Tajkhorshid et al (unpublished)

Mutant Conformational Changes

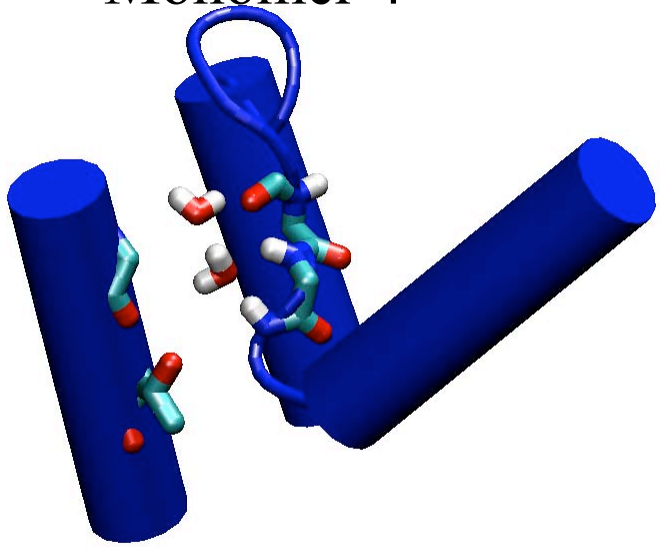


Before mutation
(t = 5 ns)
all monomers

Monomer 3



Monomer 4

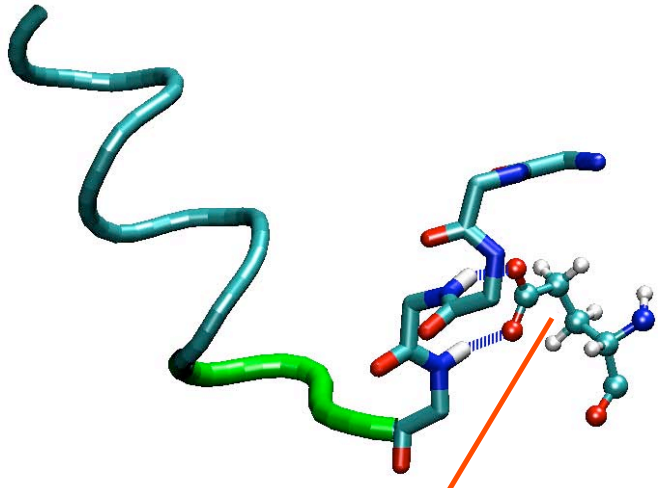


t = 20 ns

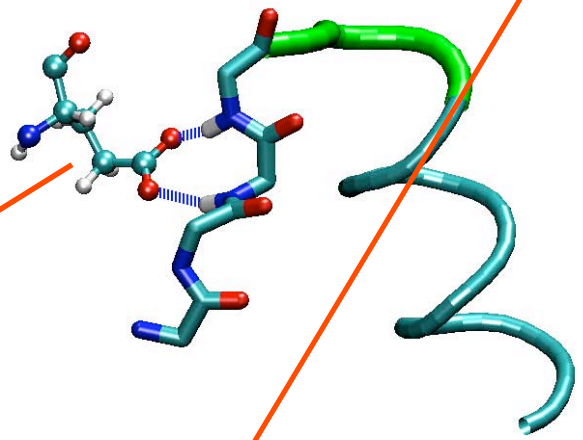
E. Tajkhorshid et al (unpublished)

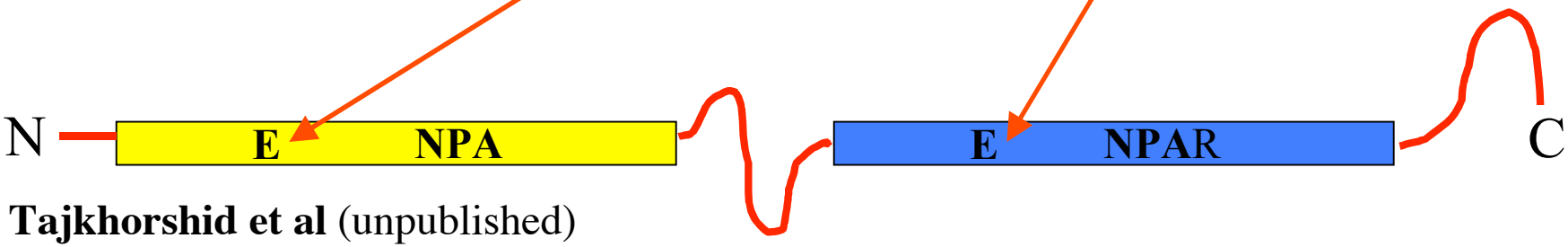
Stability of the non-helical parts

The glutamate residues stabilize the secondary structure of the inverted helices



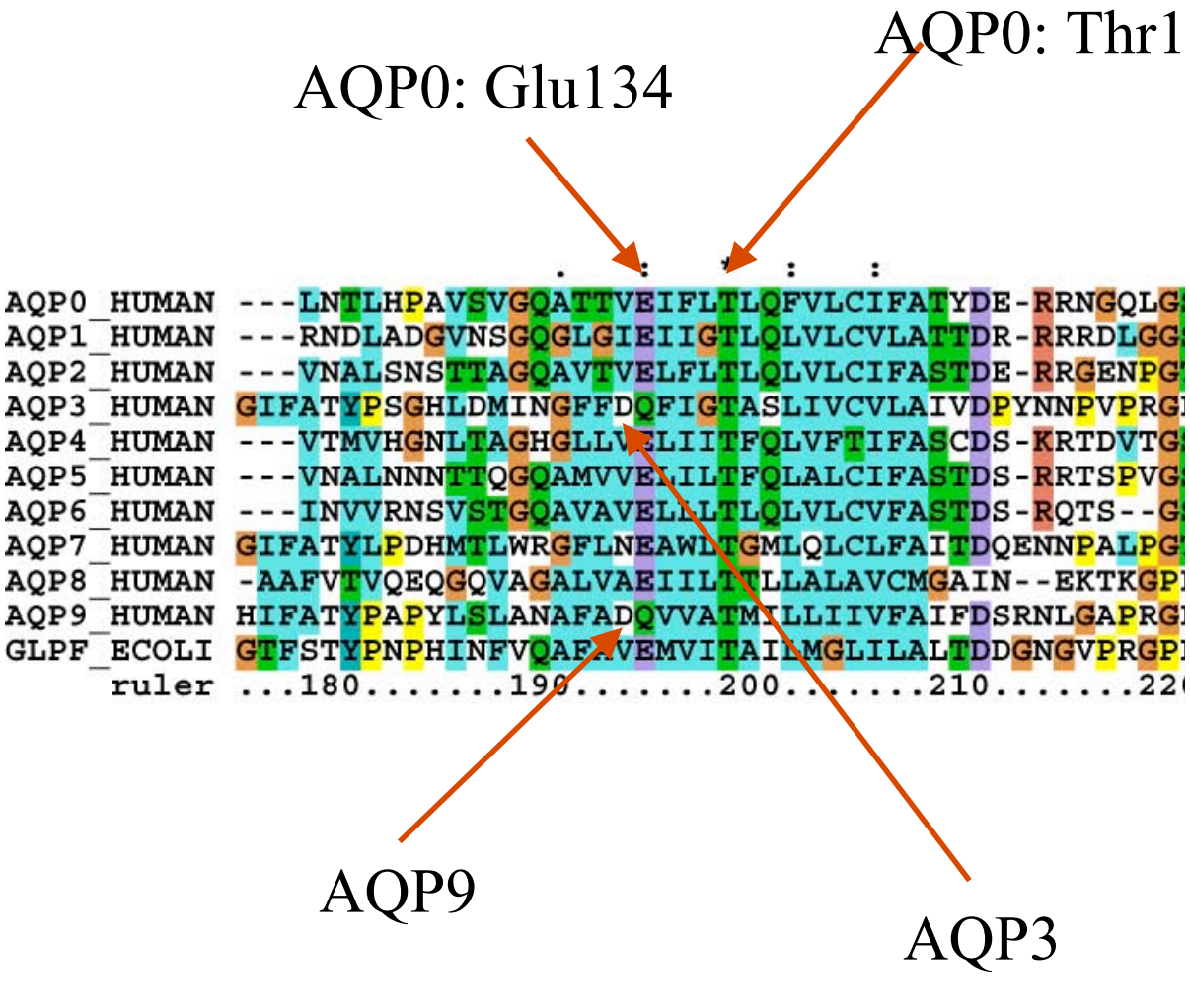
The only glutamate residues buried in the transmembrane region of the channel





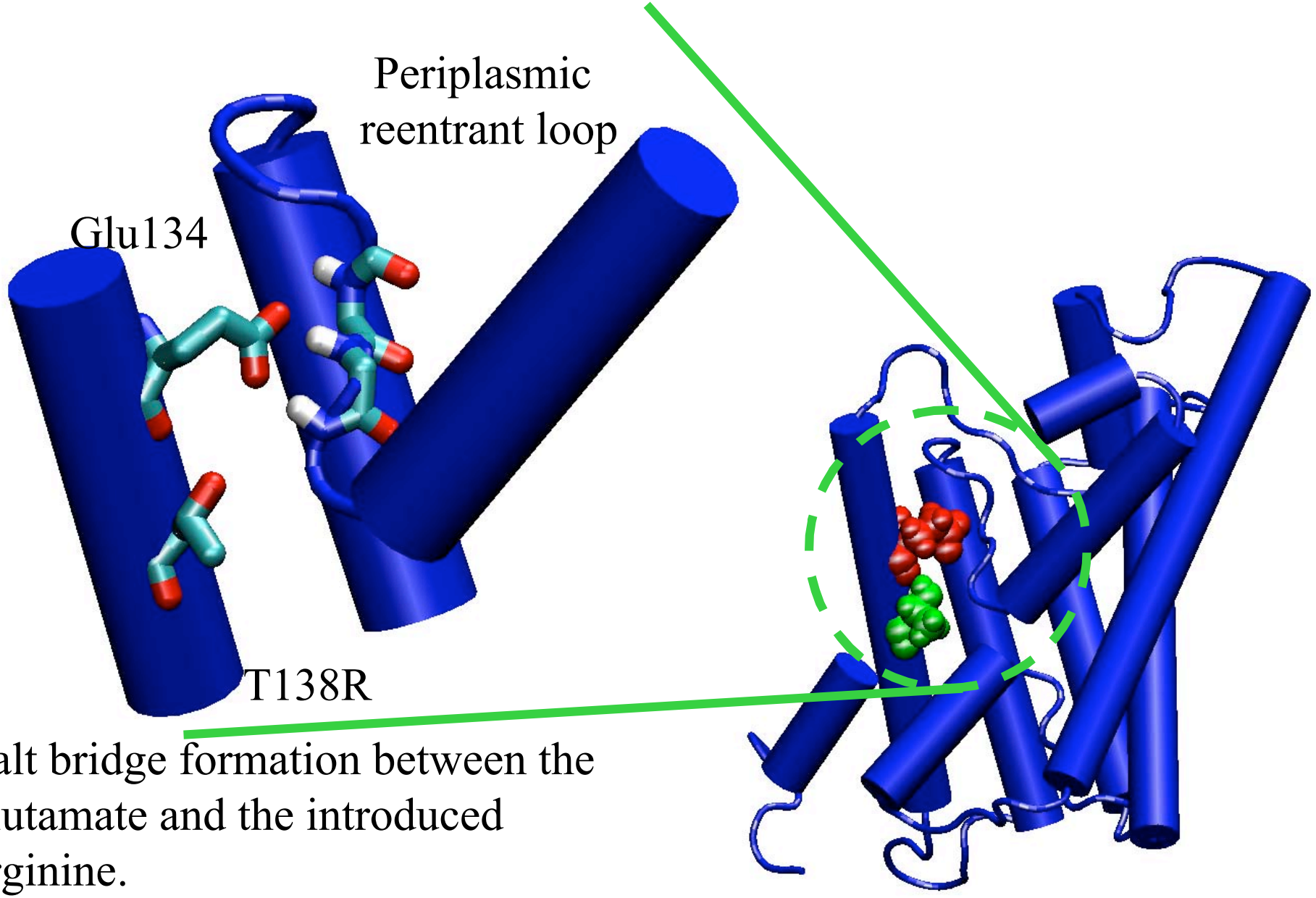
E. Tajkhorshid et al (unpublished)

Behaviour of Wild Type Aquaporins



Whenever GLU is missing, an ASP is present

Plausible Mechanism for T138R mutation

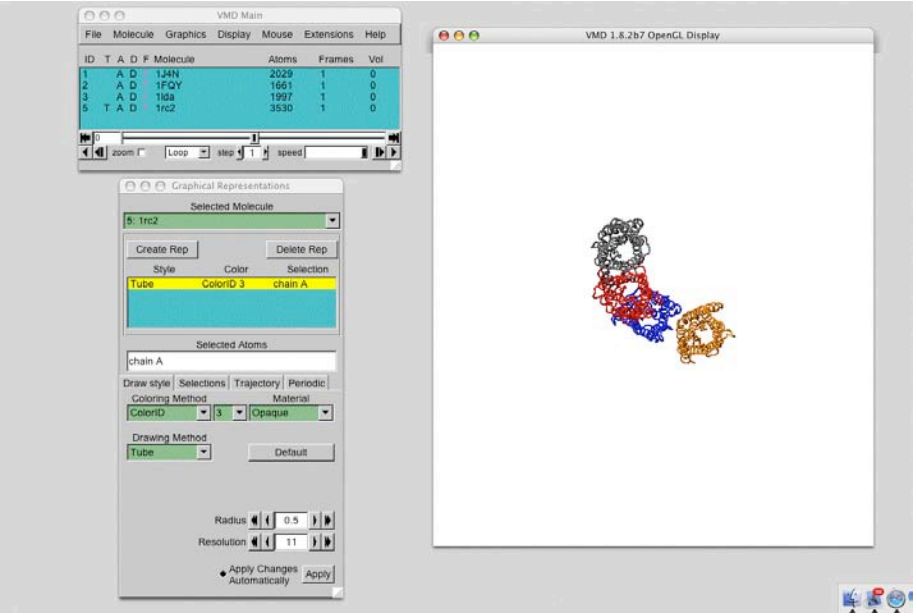


Salt bridge formation between the glutamate and the introduced arginine.

Tutorial

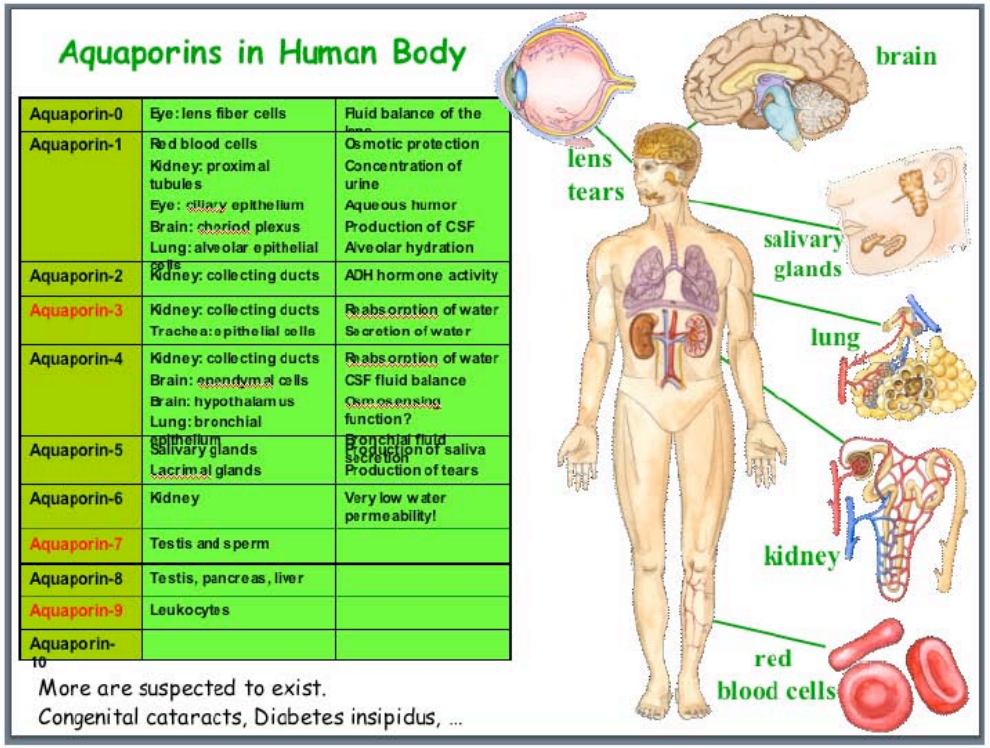
will be available at
www.ks.uiuc.edu

after release of VMD 1.8.3
July 2004



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

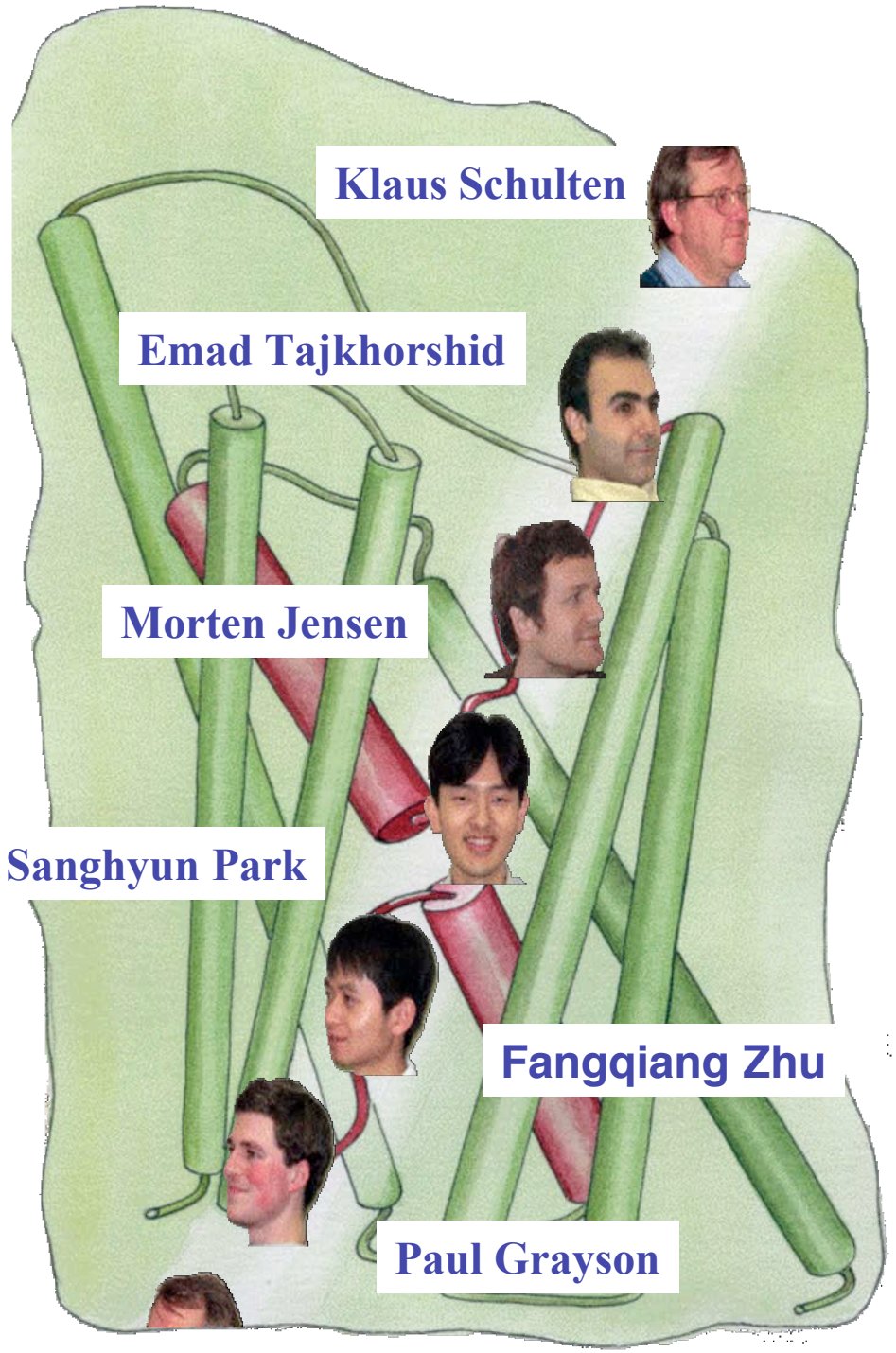
Aquaporins



Aquaporin-0	Eye: lens fiber cells	Fluid balance of the
Aquaporin-1	Red blood cells Kidney: proximal tubules Eye: ciliary epithelium Brain: choriod 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 Osmoreceptor 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		

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

National Center for Research Resources
NIH