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

AQP cluster

GLP cluster

| AQP0_HUMAN | LNTLHPAVSVGATIVEIFLTLQFLVCIFATYDE-RRNGQLGVALAVGFLALGHLFGMYTTGAGM 183 |
| AQP1_HUMAN | RNDLADGVSCNGGLBIEIIGTLQQLVCVLATDTR-RRRLLGSAPLAIGLQVALGHLLADYTTCGI 191 |
| AQP2_HUMAN | VNALSNSTTAGAVVELFLTQVLCLICFASTDE-RRQENPGRFAALSGVALGHLHGLIHYTGCSM 183 |
| AQP3_HUMAN | GFAITPSCGMNFFDDQIGTAELVICLAVLDYPNNPGVPRGLEAFTVGLVLVIGTSMGPNXGYAV 214 |
| AQP4_HUMAN | VTMVHGLTGHGLMLITIFQQLVFIFASCDSRRTDVGGTSGIALAIIGFSAIGHLFPAIYTGCSM 121 |
| AQP5_HUMAN | VNALNNNTQCQAMVEILITFQOLALCIFASTDS-RRTSPGSAISSLGVTGLHLVGYITGCSM 184 |
| AQP6_HUMAN | INVNRVSTQAVVSELCTLQVLCVFADTS-QRTS-GSPTMGISWALGHLILIGLFTGCSM 195 |
| AQP7_HUMAN | GFAITYPDPHTLNLWGFLNEAHLGMLQCLFAYITQENPALPETELVIGILVIIIGVLSLMDTYAI 225 |
| AQP8_HUMAN | AAFVTVQEQQVAGLVAEIIITLTVLALAVCMAIN--EKTQGAPFSIGFAVTVDILACGPNVSGCM 209 |
| AQP9_HUMAN | HIFATYPAHYLSNADQVATMLILITVFAIFDSRNLRAPRGLPQIAGLLTVIALSSGSLGNSGCM 215 |
| GLPF_ECOLI | GTSFTYPNPINFQAAAVEMVTAIIMGILALTDDNGVPRGPLAPLGLIAGVAGSMQFGTGFAM 202 |

ruler ...180........190........200........210........220........230........240.....
The Aquaporin Superfamily

# Water and Glycerol Channels in the Human Body

<table>
<thead>
<tr>
<th>Aquaporin</th>
<th>Location/Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aquaporin-0</td>
<td>Eye: lens fiber cells Fluid balance of the lens</td>
</tr>
<tr>
<td>Aquaporin-1</td>
<td>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</td>
</tr>
<tr>
<td>Aquaporin-2</td>
<td>Kidney: collecting ducts ADH hormone activity</td>
</tr>
<tr>
<td>Aquaporin-3</td>
<td>Kidney: collecting ducts Trachea: epithelial cells Reabsorption of water Secretion of water</td>
</tr>
<tr>
<td>Aquaporin-5</td>
<td>Salivary glands Lacrimal glands Production of saliva Production of tears</td>
</tr>
<tr>
<td>Aquaporin-6</td>
<td>Kidney Very low water permeability!</td>
</tr>
<tr>
<td>Aquaporin-7</td>
<td>Testis and sperm</td>
</tr>
<tr>
<td>Aquaporin-8</td>
<td>Testis, pancreas, liver</td>
</tr>
<tr>
<td>Aquaporin-9</td>
<td>Leukocytes</td>
</tr>
</tbody>
</table>

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)
- 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
- Cartoon
- Tube
- Tube - Transparent
Highlighting Key Conserved Residues
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

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


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

Resulting Potential of Mean Force

\[ e^{-\frac{\Delta G}{k_B T}} = \langle e^{-\frac{W}{k_B T}} \rangle \]

- The largest free energy barrier \( \approx 7.3 \text{ kcal/mol} \)

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

Simulated System 3: GlpF With Only Water

18 water molecules conducted / (4 monomers 4 ns) ➞ 1.125 water/monomer ns
Water Positions Determined by Simulations and Experiments

Electrostatic Stabilization of Water Bipolar Arrangement

Water flux vs. pressure difference

- Calculated $p_f$: $(7.1 \pm 0.9) \times 10^{-14}$ cm$^3$/s
- Experimental $p_f$ values: $5 \sim 11 \times 10^{-14}$ cm$^3$/s

Proton Exclusion in Aquaporin Channel

Initial condition: $\text{H}_3\text{O}^+$ in center

Energy barrier preventing proton conduction


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.

AQP0: Glu134
AQP0: Thr138

AQP1: Glu142
AQP1: Thr146

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

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

E. Tajkhorshid et al (unpublished)
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. Tajkhorsid et al (unpublished)
Wildtype Conformational Changes

Before mutation
(t = 5 ns)
all monomers

E. Tajkhorshid et al (unpublished)
Mutant Conformational Changes

Before mutation
(t = 5 ns)
all monomers

E. Tajkhorshid et al (unpublished)

Monomer 3

Monomer 4

(t = 20 ns)
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)
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
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