Boundary-Integral Methods in Molecular Science and Engineering

Lecture 1: Biology is Awesome.

Jaydeep P. Bardhan

Rush University Medical Center

jaydeep_bardhan@rush.edu

Outline for Lectures

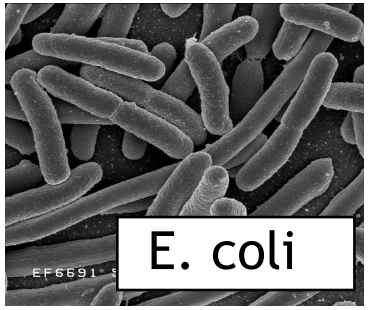
1. Biology is awesome. If you can solve Poisson, you can join in the fun!

2. There's more than one way to skin a cat. Sometimes PDEs can be advantageously reframed as *integral equations*.

3. Numerical solution of integral equations presents different challenges than PDEs.

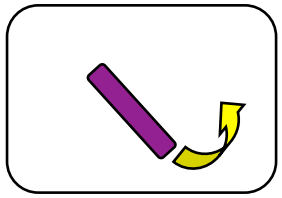
4. A diversity of unusual computational challenges will continue to drive biological simulation.

Example 1: Bacterial Chemotaxis



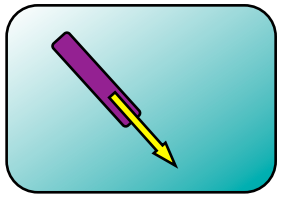
- "Nutrient receptors" cover cell membrane and guide cell towards higher-concentration areas

No food gradient

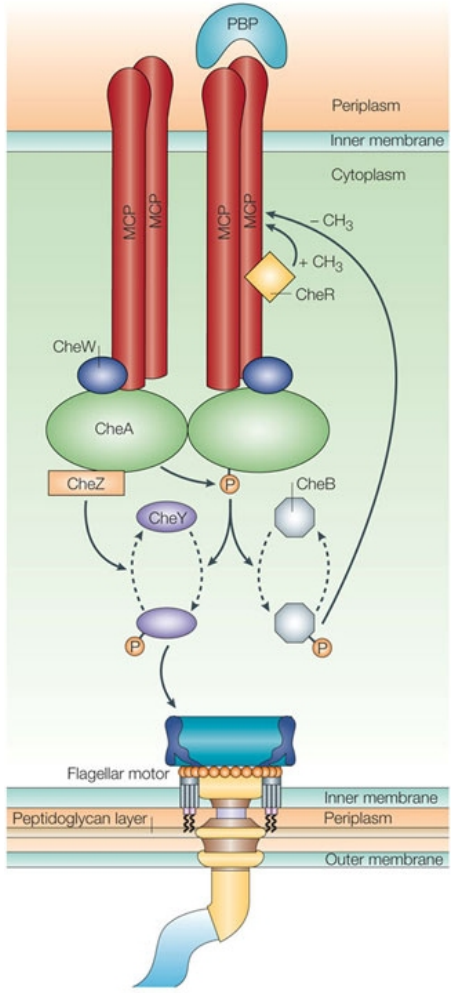


Mostly "tumbling"

Food gradient



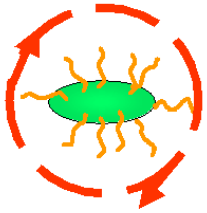
Mostly "gradient following"



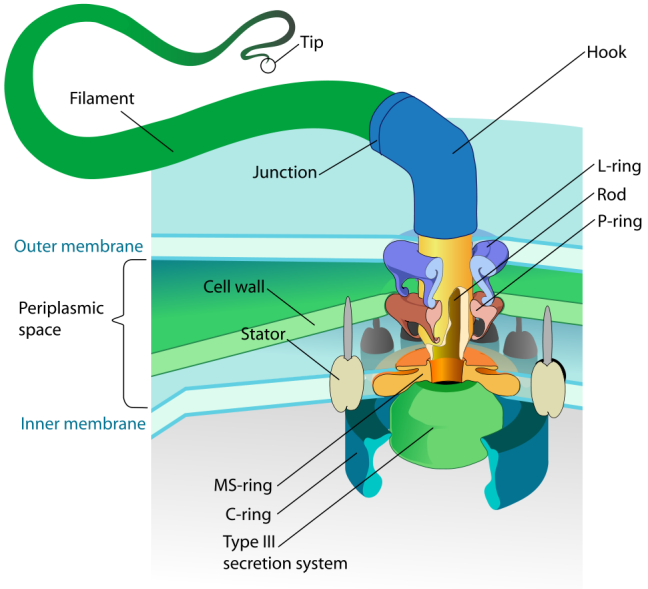
Nature Reviews | Molecular Cell Biology

Correlation of swimming behaviour and flagellar rotation in *E. coli*

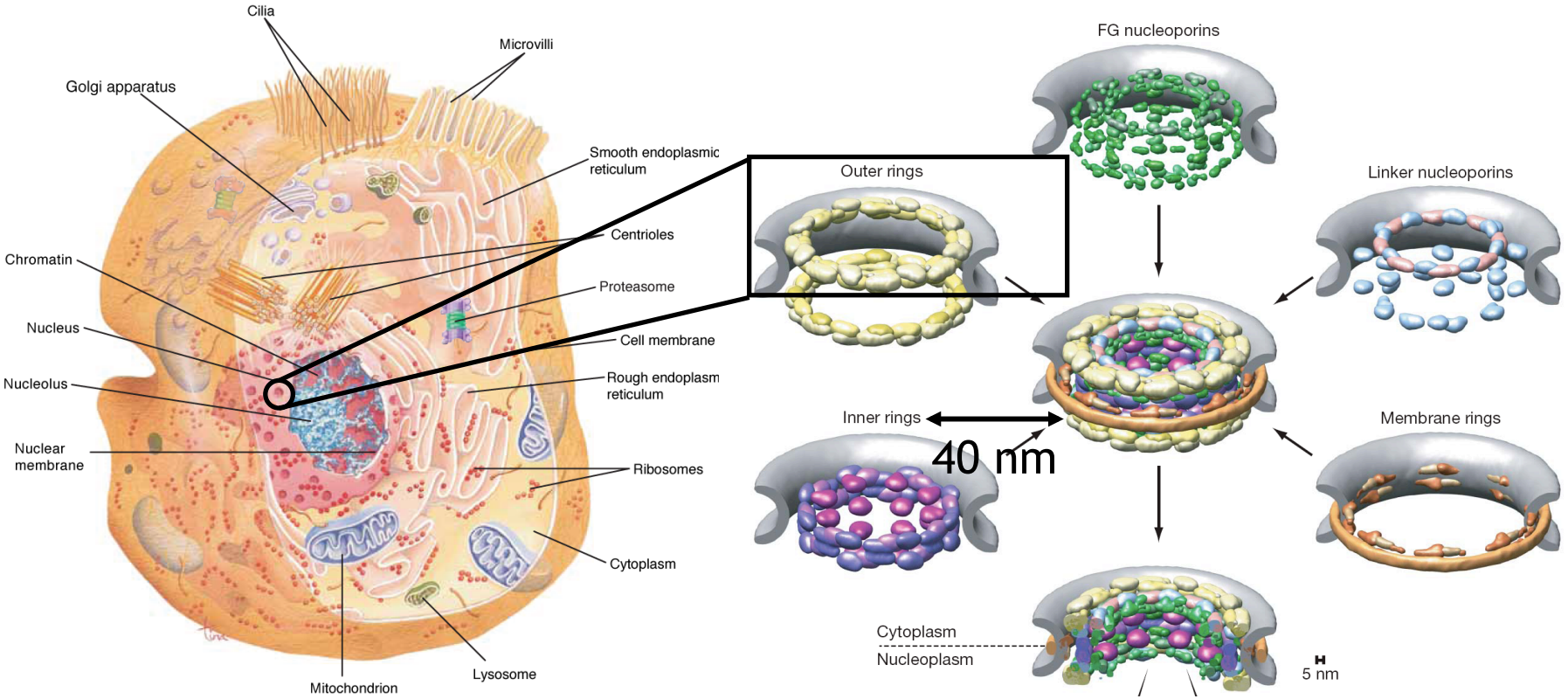
straight swim



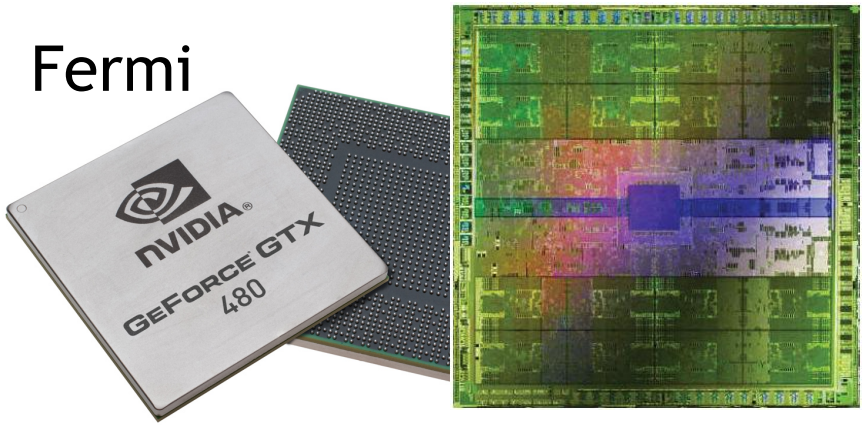
tumbling

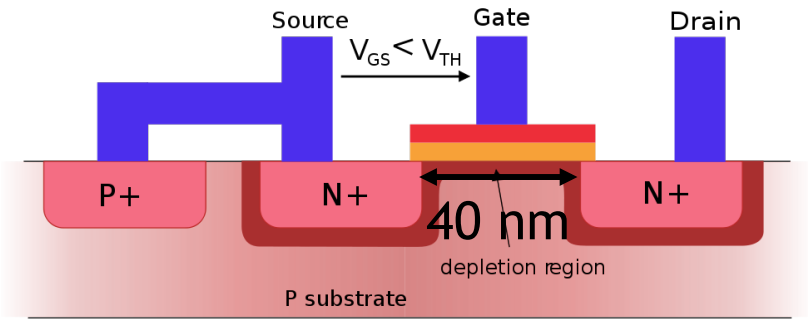


Example 2: Nuclear Pore Complex

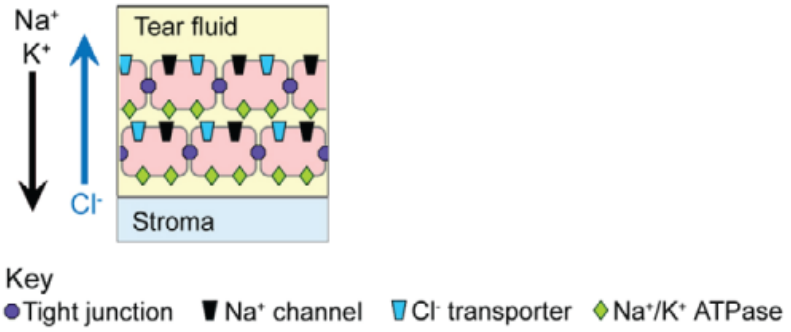


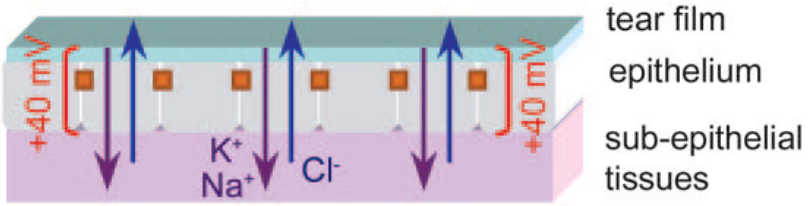
Fermi





Example 3: Electrostatics In Wound Healing





Galvani's "Figure 1"

Electric-field induced extra limb!

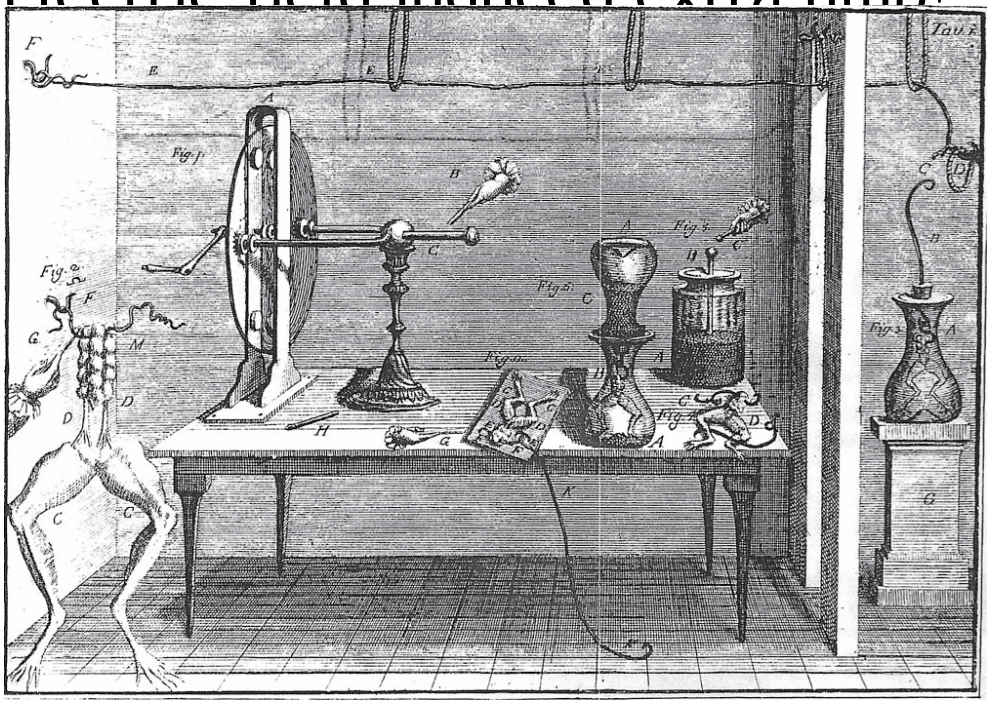
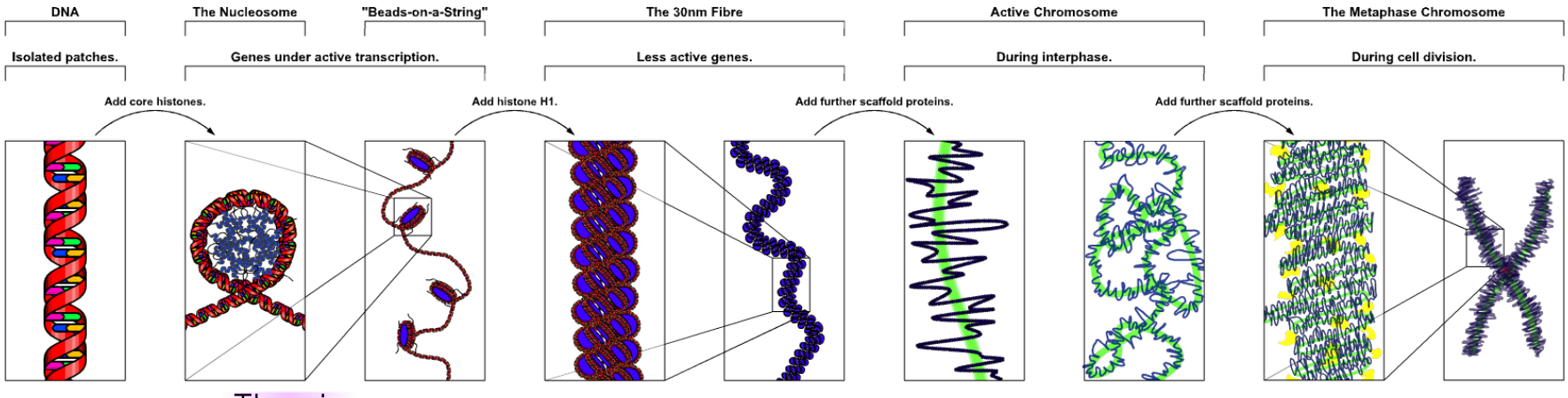
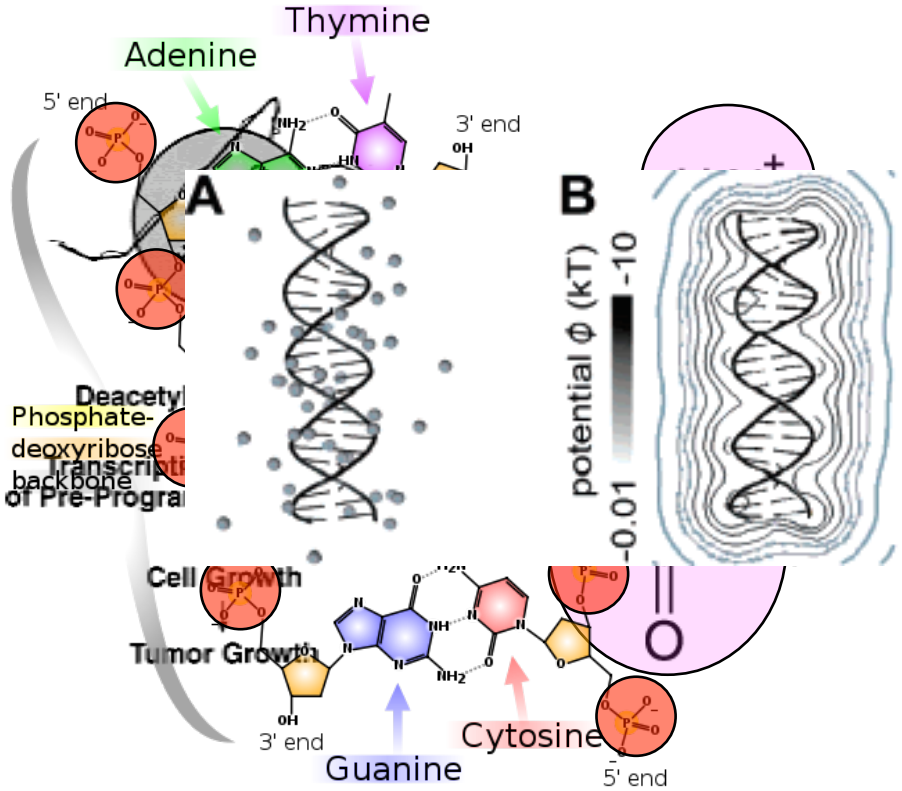
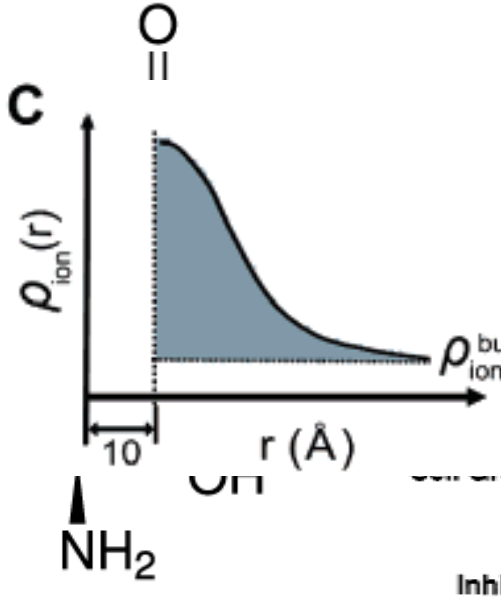


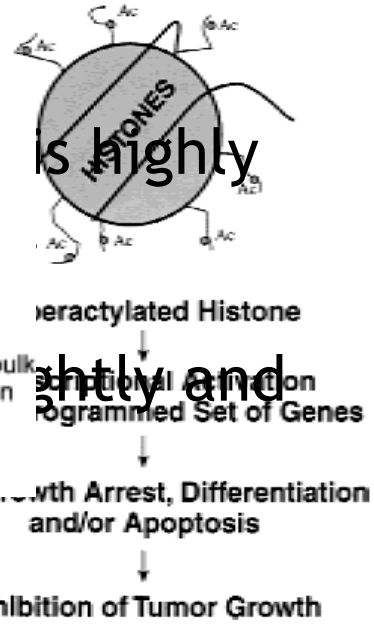
FIG. 2. Plate I of the *Commentarius* (1791 edition). The prepared frog and the electric machine on the left allude to the spark experiment.

Example 4: Compaction of DNA





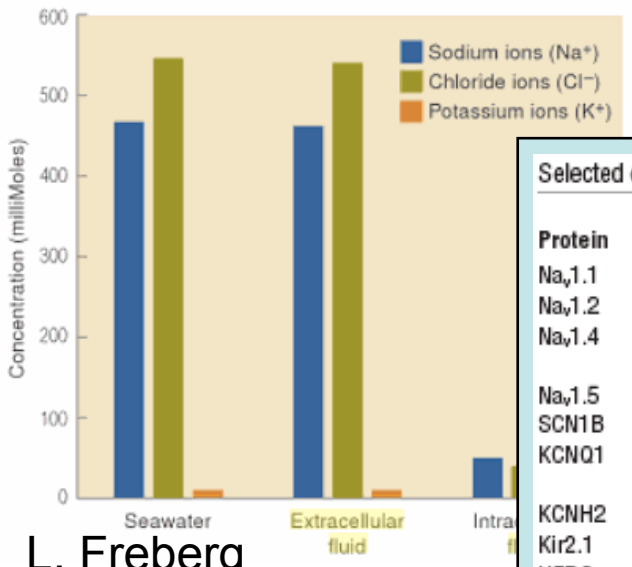




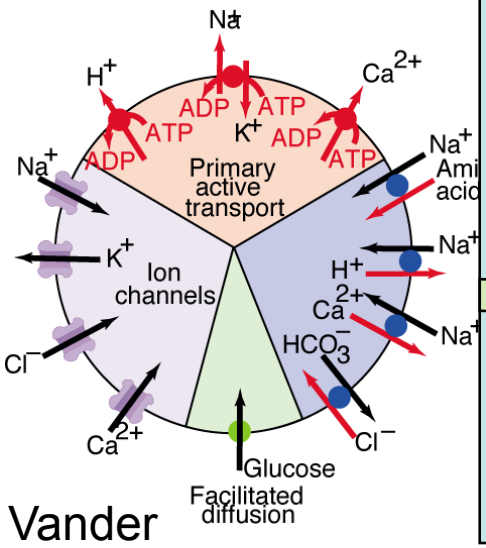
is highly

ghtly and

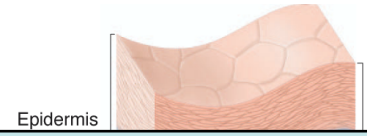
Fact: Water Makes Life Possible



L. Freberg



Vander



Selected channelopathies reviewed in this series

Protein	Gene	Disease	Functional defect
$Na_v1.1$	*SCN1A*	Generalized epilepsy with febrile seizures plus (GEFS+)	Hyperexcitability
$Na_v1.2$	*SCN2A*	Generalized epilepsy with febrile and afebrile seizures	Hyperexcitability
$Na_v1.4$	*SCN4A*	Paramyotonia congenita, potassium-aggravated myotonia, hyperkalemic periodic paralysis	Hyperexcitability
$Na_v1.5$	*SCN5A*	LQTS/Brugada syndrome	Heart action potential
SCN1B	*SCN1B*	Generalized epilepsy with febrile seizures plus (GEFS+)	Hyperexcitability
KCNQ1	*KCNQ1*	Autosomal-dominant LQTS with deafness	Heart action potential/inner ear K^+ secretion
		Autosomal-recessive LQTS	Heart action potential
KCNH2	*KCNH2*	LQTS	Heart action potential
Kir2.1	*KCNJ2*	LQTS with dysmorphic features	Heart action potential
HERG	*KCNH2*	Congenital and acquired LQTS	Heart action potential and excessive responses to drugs
Ankyrin-B	*ANKB*	LQTS	Heart action potential
$Ca_v1.2$	*CACNA2*	Timothy syndrome	Multisystem disorders
Kir6.2	*KCNJ11*	Persistent hyperinsulinemic hypoglycemia of infancy	Insulin hypersecretion
		Diabetes mellitus	Insulin hyposecretion
SUR1	*SUR1*	Persistent hyperinsulinemic hypoglycemia of infancy	Insulin hyposecretion
SUR2	*SUR2*	Dilated cardiomyopathy	Metabolic signaling
KCNE1	*KCNE1*	Autosomal-dominant LQTS with deafness	Heart action potential
		Autosomal-dominant LQTS	Heart action potential
KCNE2	*KCNE2*	LQTS	Heart action potential
CFTR	*ABCC7*	Cystic fibrosis	Epithelial transport defect
ClC-1	*CLCN1*	Myotonia (autosomal-recessive or -dominant)	Defective muscle repolarization
ClC-5	*CLCN5*	Dent disease	Defective endosome acidification
ClC-7	*CLCN7*	Osteopetrosis (recessive or dominant)	Defective bone resorption
ClC-Kb	*CLCNKB*	Bartter syndrome type III	Renal salt loss
RyR1	*RyR1*	Central core disease, malignant hyperthermia	Abnormal muscle activity
RyR2	*RyR2*	Catecholaminergic polymorphic tachycardia	Exercise-related cardiac arrhythmias

Kass ‘05

Biology's Multiscale Challenges

Modeling and Simulations

- Fast computers and MD are not a panacea...
  - The inverse problem: what kinds of molecules have property X?
  - Combinatorial number of problems to simulate

state-of-art | desired range for simulations

10^{-15} s 10^{-12} 10^{-9} 10^{-6} 10^{-3} 10^{0} s

Acetylcholinesterase | Lactate dehydrogenase

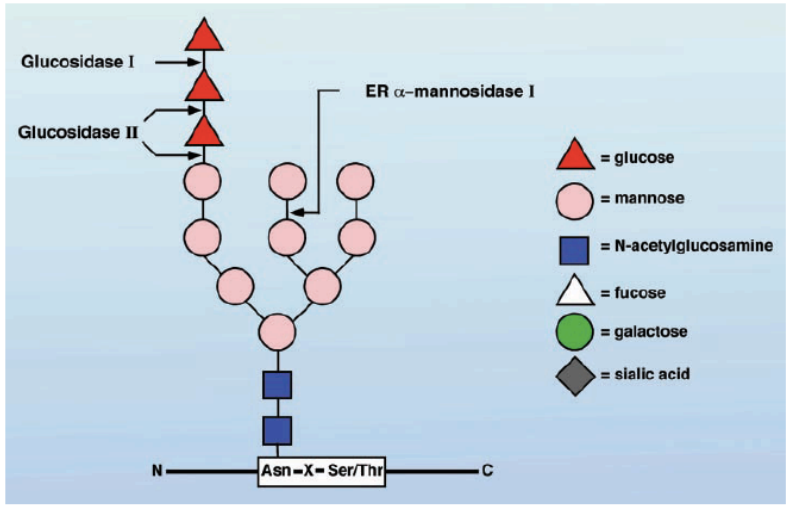


Fig. 1. The N-linked core oligosaccharide. N-linked glycans are added to proteins in the ER as "core oligosaccharides" that have the structure shown. These are bound to the polypeptide chain through an N-glycosidic bond with the side chain of an asparagine that is part of the Asn-X-Ser/Thr consensus sequence. Terminal glucose and mannose residues are removed in the ER by glucosidases and mannosidases. The symbols for the different sugars are used in the following figures.

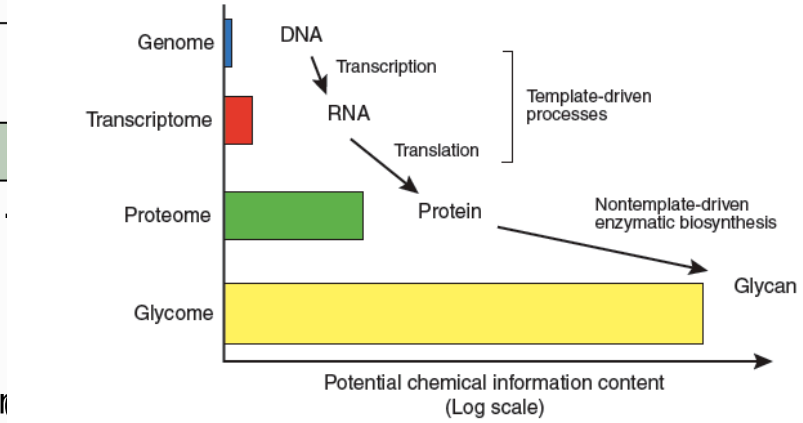


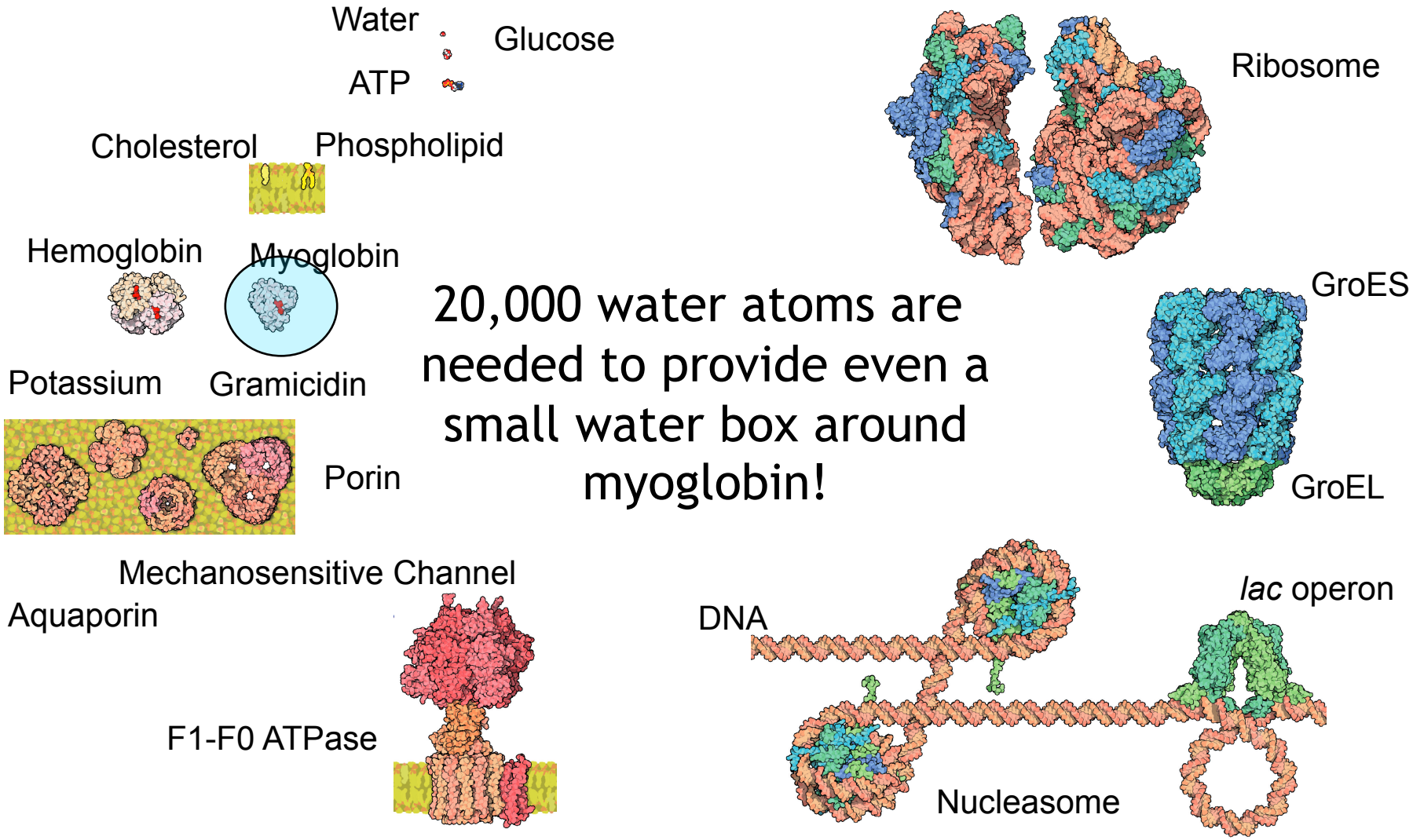
Figure 1 Glycome enhancement of the molecular and functional diversity of the proteome. Protein expression is based on a genetically encoded template, but post-translational modifications of proteins dramatically enhance their functional diversity. The glycome represents the main class of post-translational modifications, providing biological access to vast information space at minimum genetic cost. Note that in the information flow from the genome to the glycome, the biosynthesis of glycans is not encoded via a template-driven system. Note also the log scale of potential chemical information.

Turnbull (2007)

10^{-15} s 10^{-12} 10^{-9} 10^{-6} 10^{-3} 10^{0} s

Agarwal and Alam (2006); Ramasubramaniam (2007)

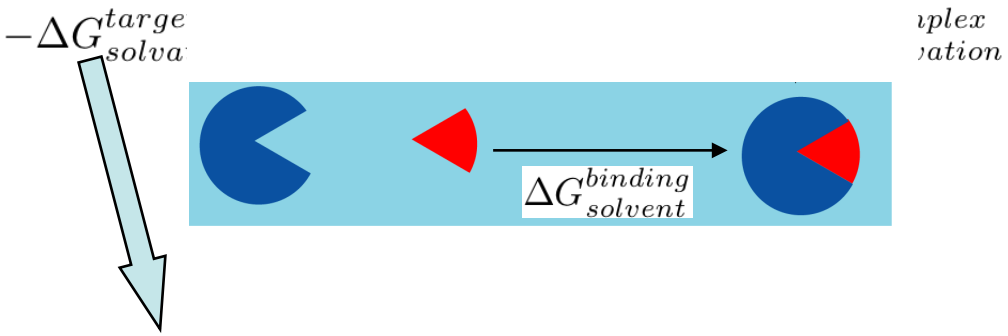
Molecular Machines, To Scale



Images by David Goodsell:
http://mgl.scripps.edu/people/goodsell

A Crucial Consequence of Solvation

- Molecular binding involves sacrificing solute--solvent interactions for solute--solute interactions:



$$\Delta G^0_{\text{solv}} = \Delta G^0_{\text{non es}} + \Delta G^0_{\text{es}}$$

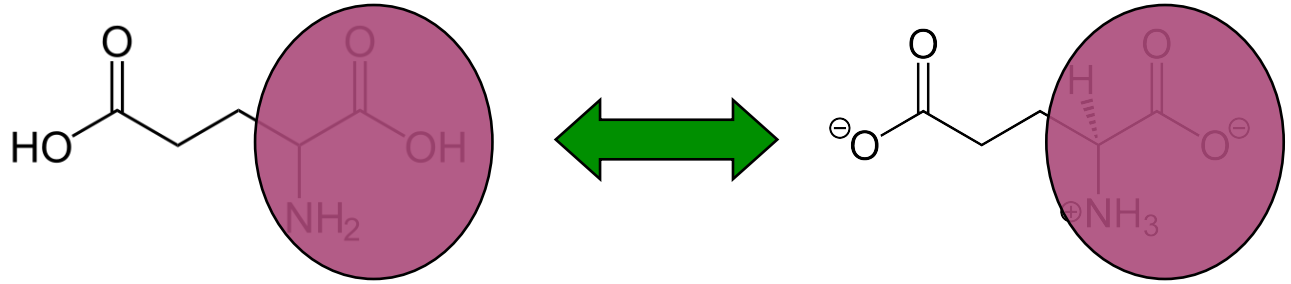
$$\Delta G^{bind} = \gamma \Delta SA + \Delta G^{es}_{L-R} - \Delta G^{es}_{L} - \Delta G^{es}_{R}$$

δ = 0 δ = 1

This is only a VERY SIMPLE MODEL for molecular binding!

Stepping Towards a Simpler Model

- To analyze protonation of ionizable groups:

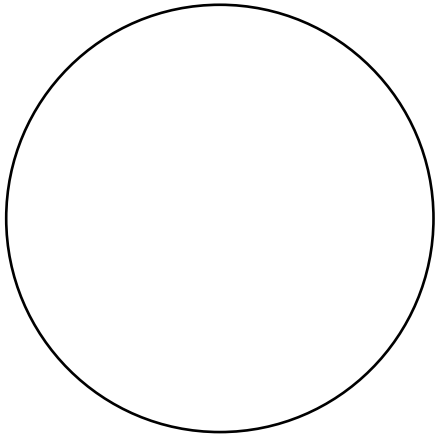


$$HA \rightleftharpoons H^+ + A^-$$

$$\epsilon_{\text{water}} \approx 80$$

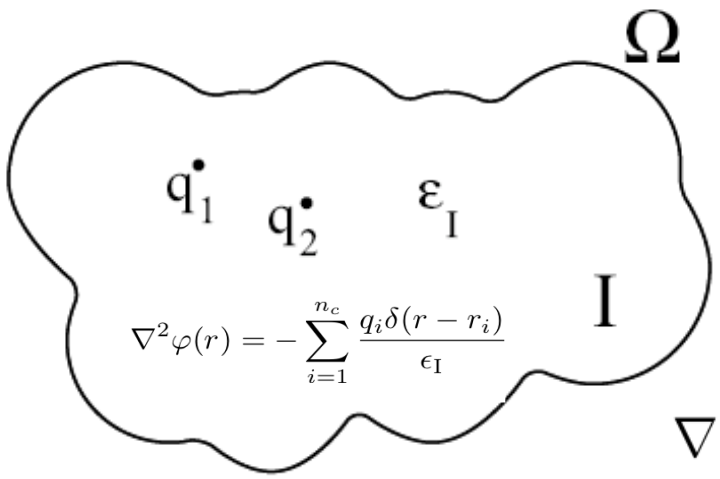
$$\epsilon_{\text{protein}} \approx 2 - 8$$

Assume spherical protein



Kirkwood, 1934; Tanford, 1957

The Basic Continuum Electrostatic Model



$$\nabla^2\varphi(r) = 0$$

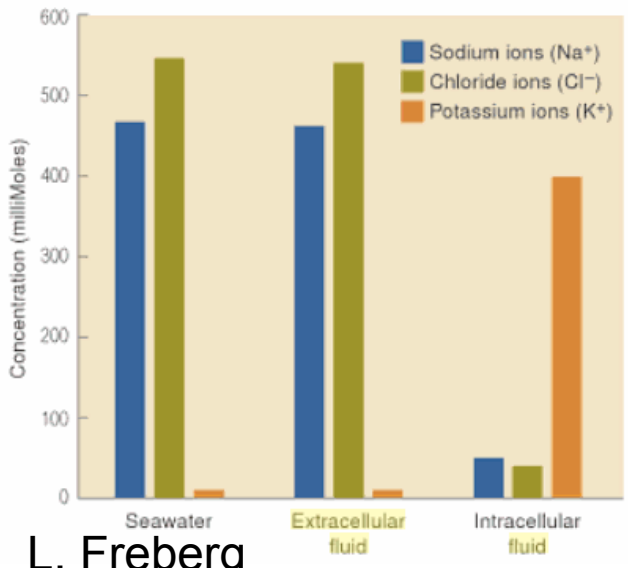
$$\Delta G = \frac{1}{2}\varphi^T q$$

- Mixed-dielectric Poisson problem with point charge sources
- Assumes molecule is at *infinite dilution*--more on that later
- Hundreds to thousands of times faster than MD: seconds to a few hours

This model can be derived rigorously from sophisticated statistical-mechanical theories (see, e.g., Beglov+Roux, 1996)

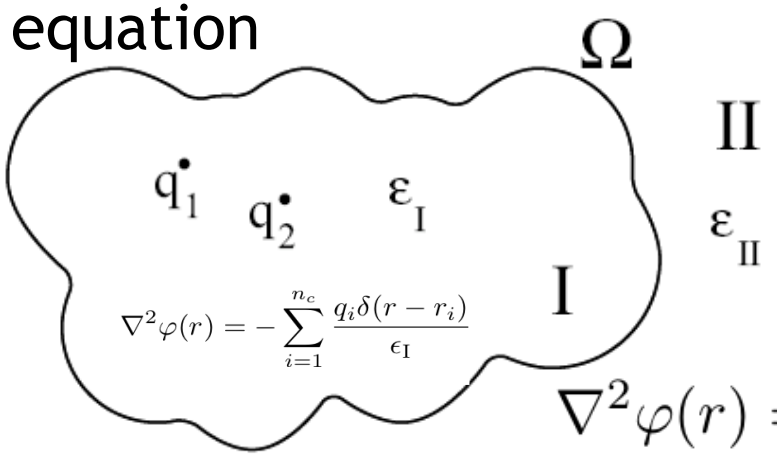
Biological Fluids are Electrolytes

- Mobile ions in solvent redistribute in response to electric fields



L. Freberg

- Assuming that the charge density in the solvent is Boltzmann-like, we obtain the Poisson-Boltzmann equation



$$\nabla^2\varphi(r) = \kappa^2\varphi(r)$$

Linearize

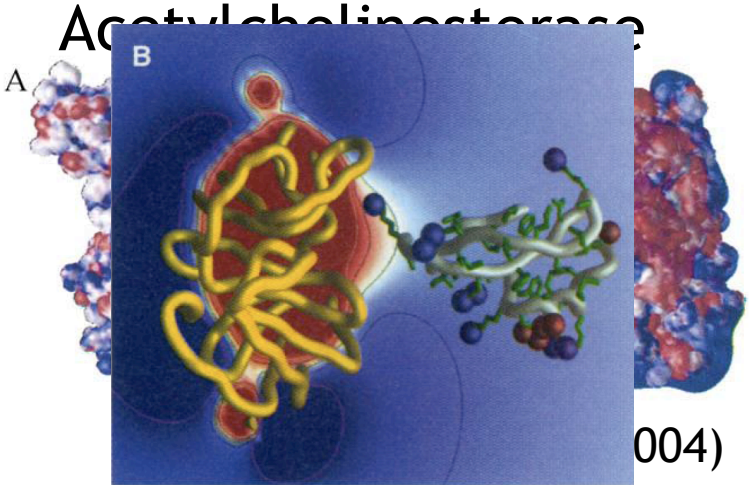
$$\nabla^2\varphi(r) = \kappa^2 \sinh(\varphi(r))$$

This treatment of ionic solutions is very useful, but incredibly oversimplified--more details in Lecture 4!

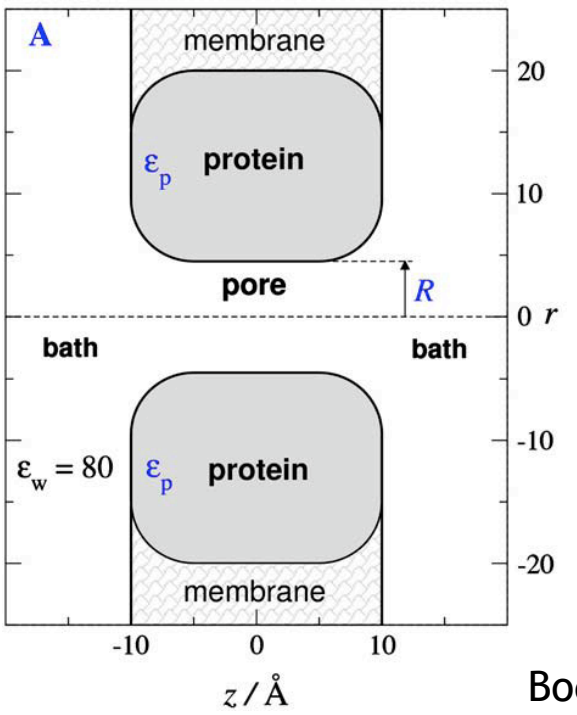
Applications of the Continuum Model

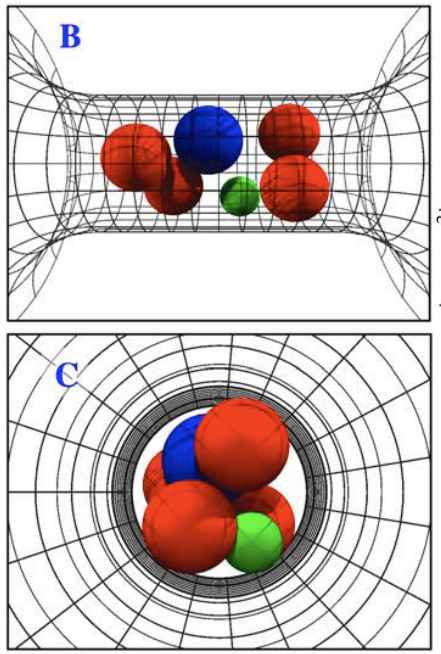
- How dielectric boundaries "focus" electric fields

- How ion channel proteins selectively pass one species but not others

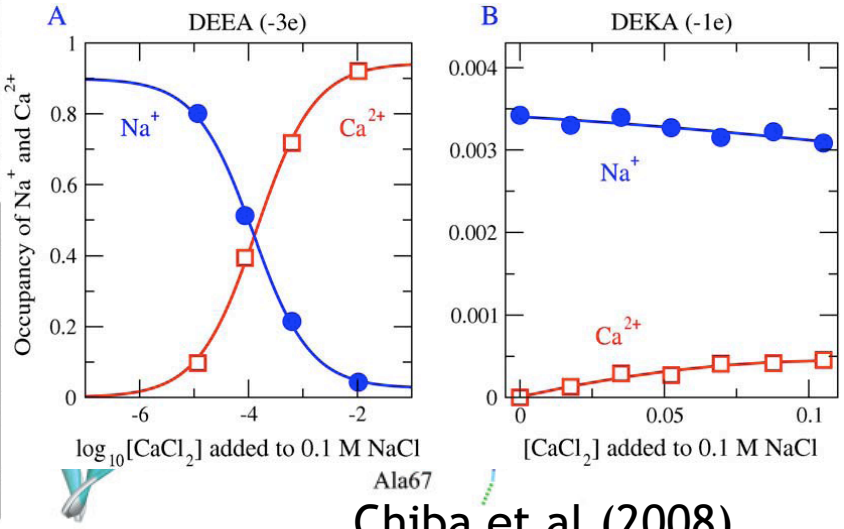


Honig+Nicholls (1995)





Boda, Gillespie, Eisenberg et al (2007)



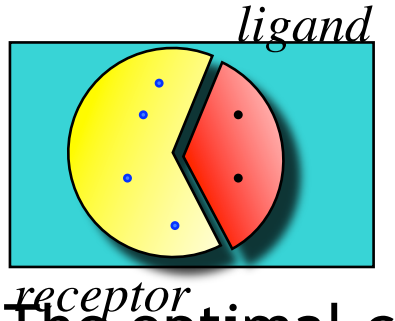
Chiba et al (2008)

Continuum Models Capture Important Physics

- Linear response means quadratic energy:

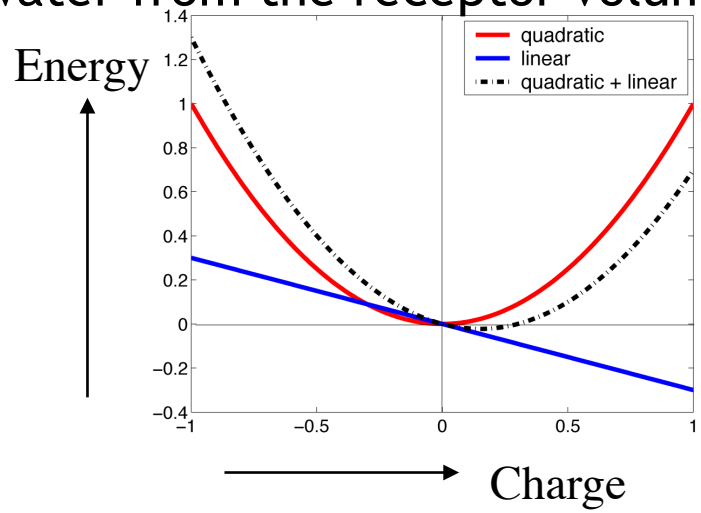
$$\varphi^{\mathrm{REAC}} = Lq \longrightarrow E = \frac{1}{2}\varphi^{\mathrm{REAC},T} q = \frac{1}{2} q^T L q$$

Assume ligand rigidity, and no charge transfer:

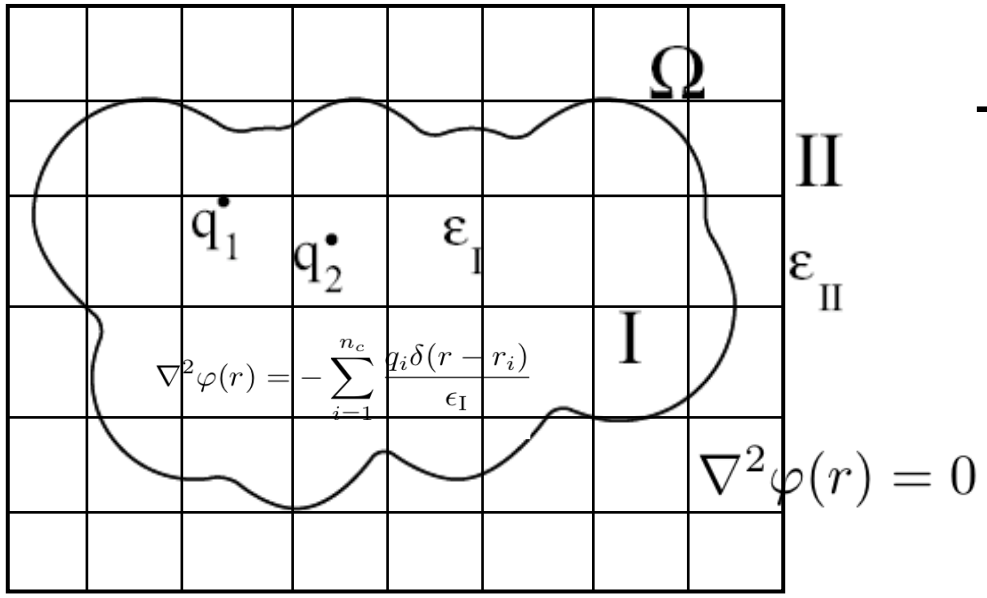


$$\Delta G^0_{\mathrm{var,es}} = \frac{1}{2} q^T \left(L_b - L_u\right) q + c^T q$$

- Under our assumptions, this energy function is convex
- The idea: It always *costs energy* to remove the water from the receptor volume
- The optimal charge distribution…
  - … balances the "desolvation penalty" against ligand-receptor interactions
  - … is a *guide* for design



Solving the PDE Directly is Possible, But…

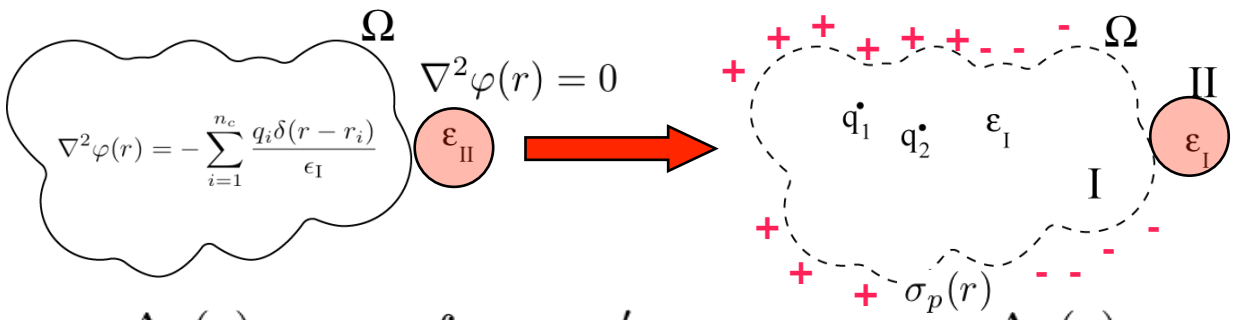


The idea: Just throw down a finite-difference grid or a finite-element mesh and go to town!

PDE Complications

1. Boundary conditions are at infinity
2. Point charges must be spread onto the grid
3. The dielectric interface is approximated

A Boundary Integral Method



$$\Delta\epsilon = \epsilon_{II} - \epsilon_I$$
$$\bar{\epsilon} = \frac{1}{2}(\epsilon_{II} + \epsilon_I)$$
$$\hat{\epsilon} = \Delta\epsilon/\bar{\epsilon}$$

$$\sigma_p(\mathbf{s}) + \frac{\Delta\epsilon(\mathbf{s})}{4\pi\bar{\epsilon}(\mathbf{s})}\mathbf{n}(\mathbf{s}) \cdot \int_\Omega \frac{\mathbf{s}-\mathbf{s}'}{|\mathbf{s}-\mathbf{s}'|^3}\sigma_p(\mathbf{s}')d\mathbf{s}' = -\frac{\Delta\epsilon(\mathbf{s})}{4\pi\bar{\epsilon}(\mathbf{s})}\mathbf{n}(\mathbf{s}) \cdot \sum_k \frac{q_k}{\epsilon(\mathbf{r}_k)} \frac{\mathbf{s}-\mathbf{r}_k}{|\mathbf{s}-\mathbf{r}_k|^3}$$

$$(I - \hat{\epsilon}\mathcal{K})\,\sigma = \hat{\epsilon} f$$

$$Ax = Bq$$

1. Boundary conditions handled exactly
2. Point charges are treated exactly
3. Meshing emphasis can be placed directly on the interface

Differential vs Integral Equations

Differential operator

$$\nabla^2 \varphi(r) = -\rho(r)$$

Unknown Source/forcing

Integral operator

$$\int_\Omega u(x) K(x;x')\,dx = f(x)$$

Unknown Source/forcing

Kernel: for our problems, this will be the Green's function of the corresponding PDE

Must specify boundary conditions ⟺ BCs automatically included (generally)

Operator is local ⟺ Operators are often **nonlocal**

Kinds of Integral Equations

First kind

$$\int_D u(x)K(x;x')dx = f(x)$$

Second kind

$$u(x) + \int_D u(x)K(x;x')dx = f(x)$$

Fredholm: domain of integration is fixed (above: D)

Volterra: domain integration varies with the independent variable

$$x(t) + \int_a^t K(t,s,x(s))ds = y(t)$$

$$x(t) = x_0 + \int_a^t f(s,x(s))ds$$

Closely related to ODEs

$$\frac{dx}{dt} = f(t,x(t)$$

$$t \geq a$$

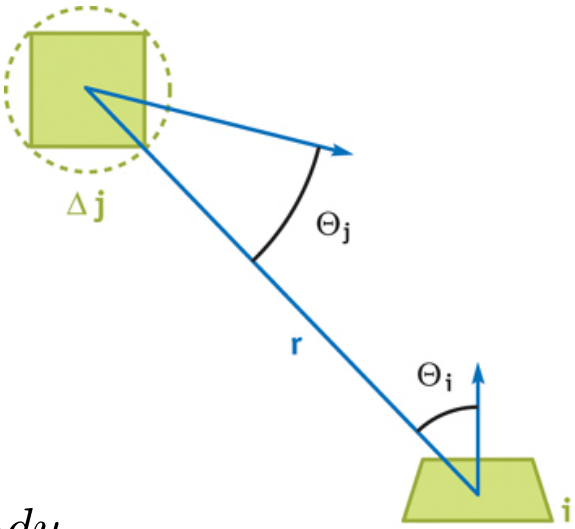
$$x(a) = x_0$$

Integral equations can be linear or nonlinear.

Applications of Integral Equations

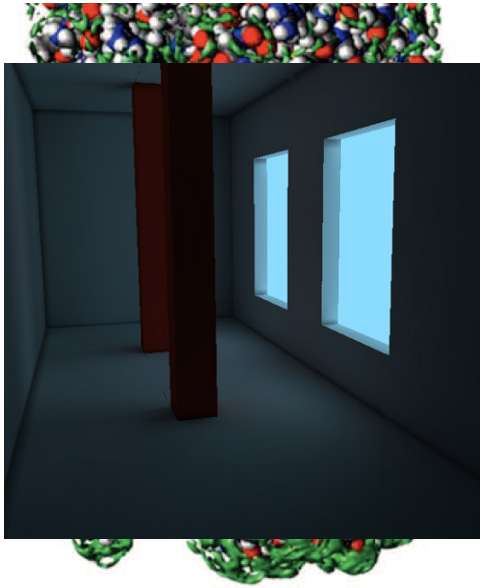
- Wiener-Hopf integral equations

$$\lambda x(t) - \int_0^\infty k(t-s)x(s)ds = y(t)$$



- Radiosity in graphics

$$B(x) = B^e(x) + \rho(x) \int_{M^2} \frac{c\cos(\theta_x)\cos(\theta_y)V(x,y)}{|x-y|^2} B(y)dy$$

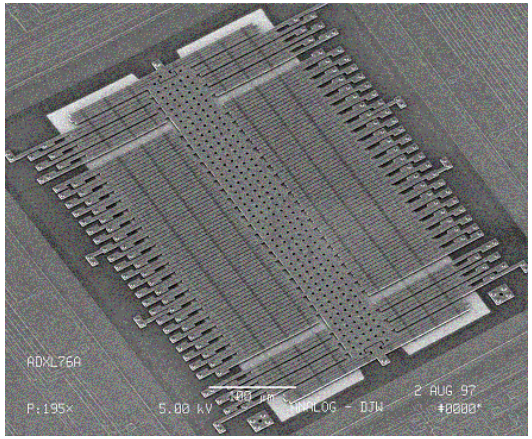


- Fluid structure in chemistry

$$\rho(r) = \rho_0 \exp\left[-\beta U(r) + \int c(r,r')\left(\rho(r') - \rho_0\right) dr'\right]$$

Boundary Integral Equations: Fluids

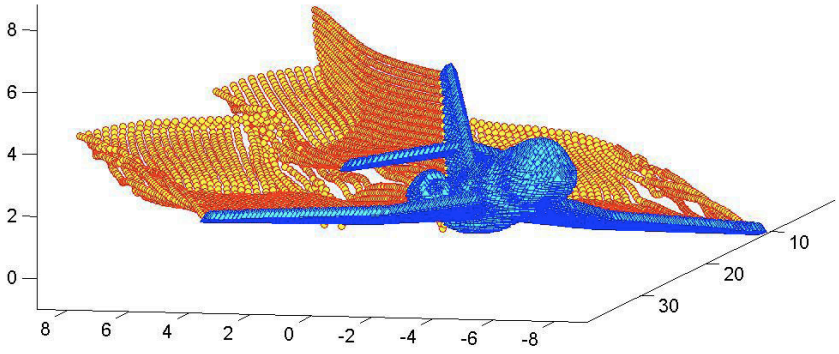
- Stokes: steady/unsteady incompressible, steady linearized compressible

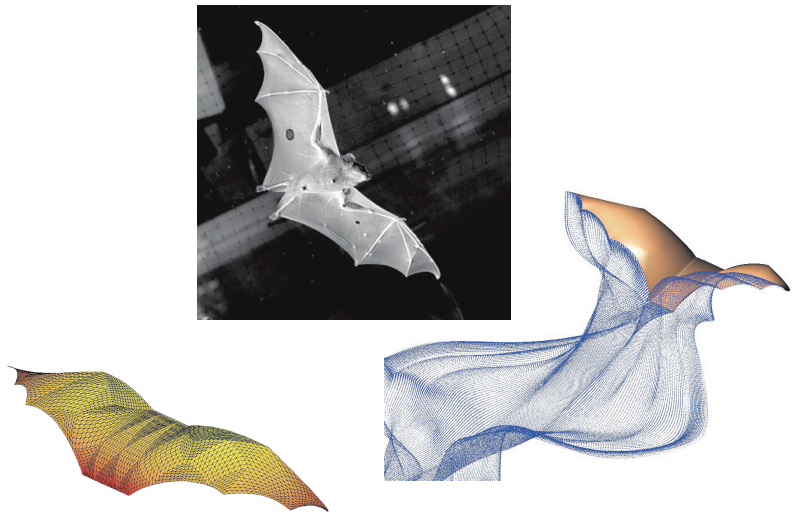


ADXL76 accelerometer

(Analog Devices; picture courtesy Xin Wang and Jacob White)

- Wide applications in potential flow

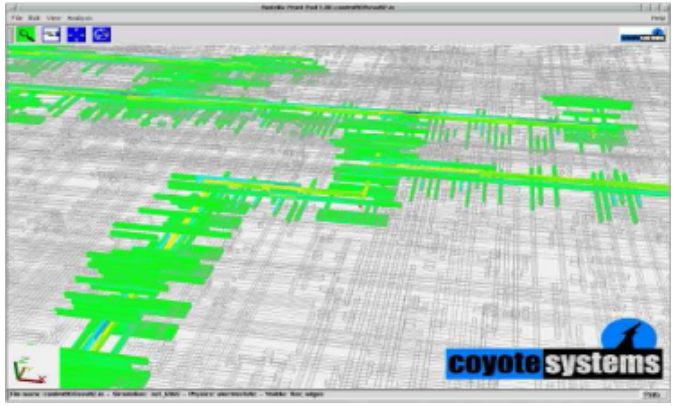




(Images courtesy David Willis)

BIE Application: Electromagnetics

VLSI interconnect analysis: extracting parasitic capacitances and inductances in complex geometries



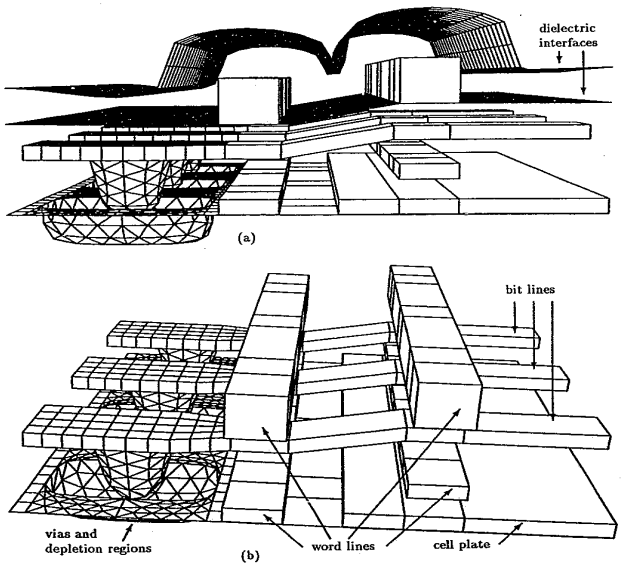
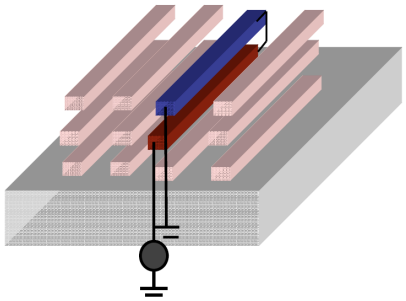
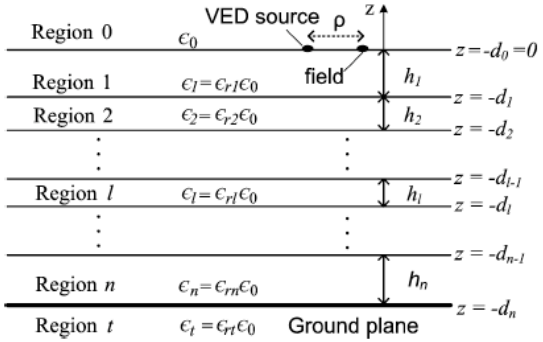


FIGURE 4-7: The complete DRAM model, (a), and with the dielectric interfaces removed for clarity, (b).

K. Nabors

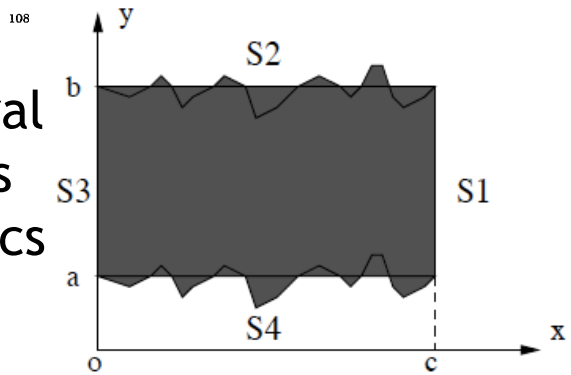
Modified Green's functions allow treatment of planar multi-laver substrates





Hu, Daniel et al (2007)

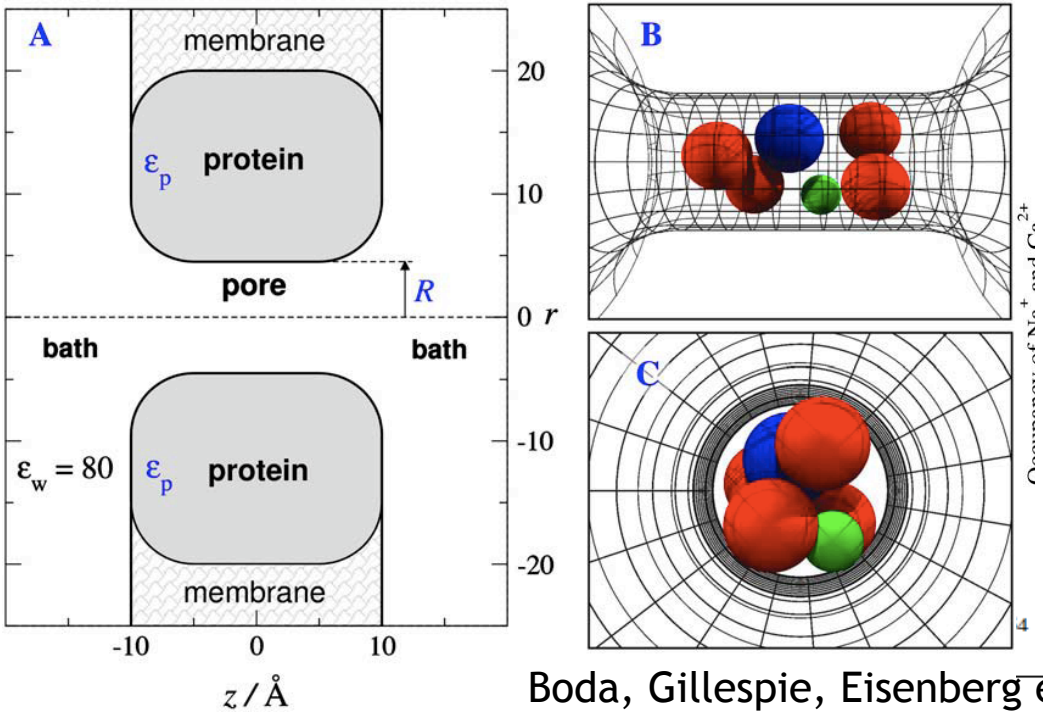
Stochastic integral equation: gives *average* parasitics



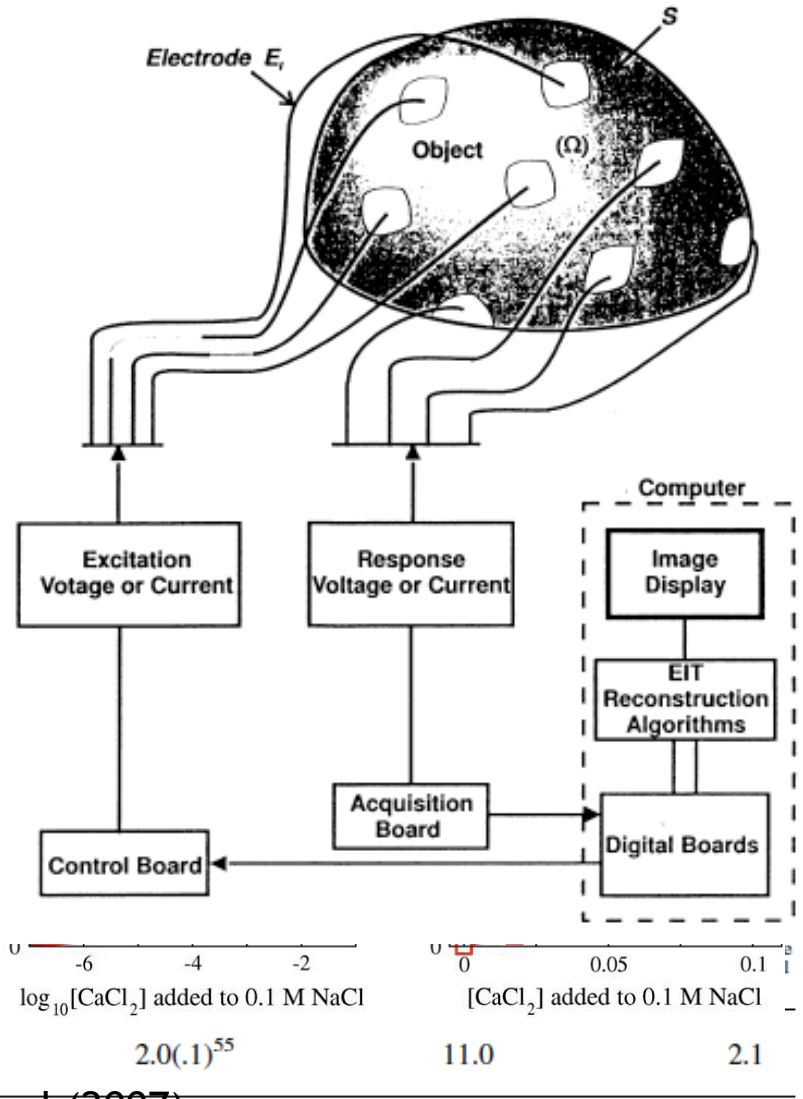
Z. Zhu (2004)

BIE Applications In Bioscience

- Protein folding
- Molecular transport
- Ion channel selectivity



Boda, Gillespie, Eisenberg et al (2007)



Duraiswamy et al (1998)

Aragon (2004)

Next: Turning PDEs into BIEs

- Advantages and disadvantages of each

- The boundary element method (BEM) for solving BIEs numerically

References:

1. Gilson, M. http://pharmacy.ucsd.edu/labs/gilson/ce_www1a.pdf

2. Sharp, K.A. and Honig, B. 1990. *Electrostatic interactions in macromolecules: Theory and applications*. Annu. Rev. Biophys. Biophys. Chem. 19: 301-332.

3. Roux, B. and Simonson T. 1999. *Implicit Solvent Models*. Biophysical Chemistry. 78:1-20.

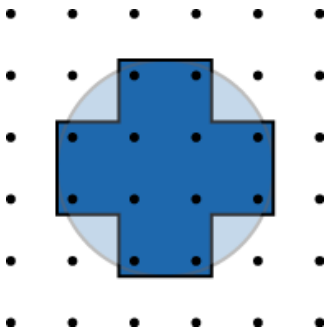
4. Bardhan, J. thesis http://www.rle.mit.edu/cpg/people_alumni.htm

5. Hildebrandt, A. thesis
   http://scidok.sulb.uni-saarland.de/volltexte/2007/1400/

Why Use Integral Equations?

Finite-Difference Method (FDM)

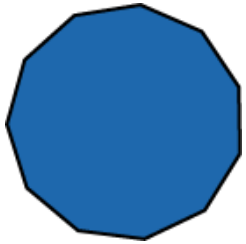
- Force discretized PDE to be satisfied at grid points



- Advantages:
  1. Can be applied to quite general problems
  2. Efficient, well-established methods exist
- Disadvantages:
  1. Grid representation introduces numerous errors
  2. Volume discretization results in poor scaling of computer resources for accaccuracy

Boundary-Element Method (BEM)

- Solve discretized integral-equation formulation



- Advantages:
  1. Surface discretization requires one dimension fewer unknowns
- Disadvantages:
  1. Discretized systems are dense
  2. Finding accurate problem representation is hard
  3. Matrix elements can be difficult to compute
  4. Complex problem geometries require complicated, problem-specific integral equations