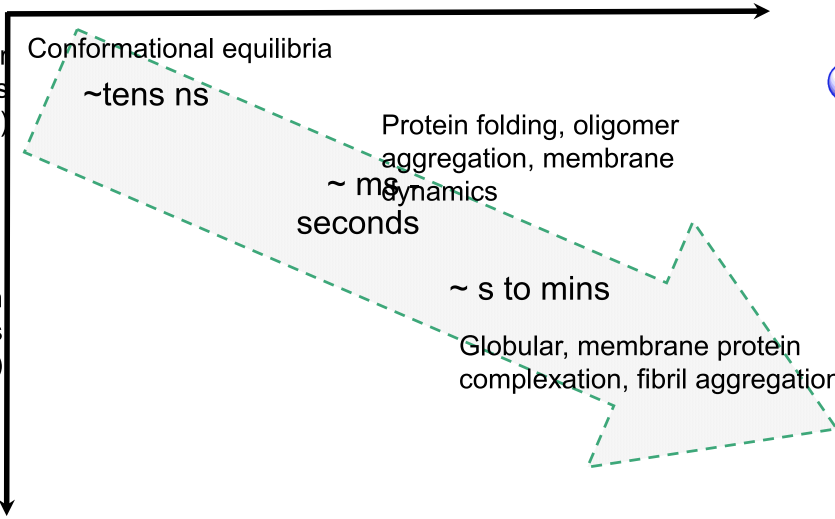
Coarse-Graining Methods in Biology and Materials

School on Multiscale Modeling and Simulations of Hard and Soft Materials

Professor Teresa Head-Gordon
Department of Bioengineering
University of California, Berkeley

Lecture 1

Coarse-Graining and MultiScale



Spatial Resolution

Atomistic
(Angstroms)

Proteins and Oligomers
(nanometers)

Protein Assemblies
(microns)

Time Resolution

Molecular
Dynamics
(nanoseconds)

Langevin
Dynamics
(microsecond)

Brownian
Dynamics
(seconds)

Conformational equilibria

~tens ns

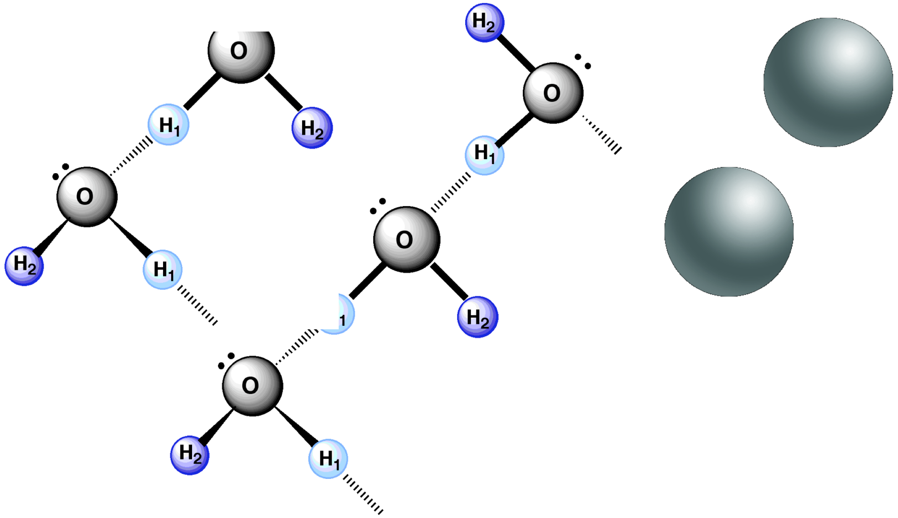
Protein folding, oligomer aggregation, membrane dynamics

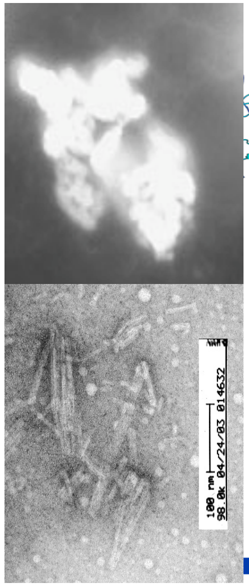
~ ms

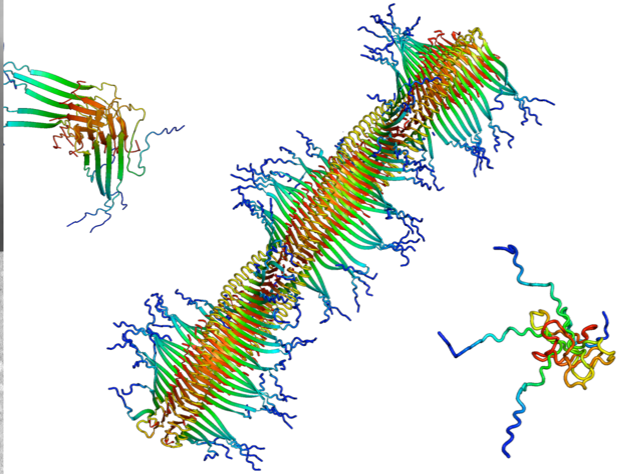
seconds

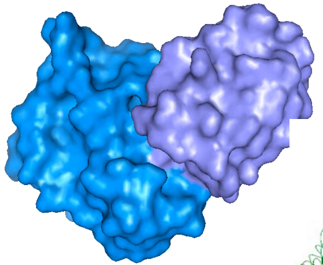
~ s to mins

Globular, membrane protein complexation, fibril aggregation

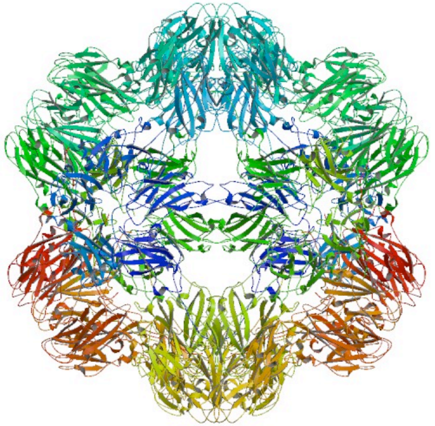








Barnase-Barstar



Brome Mosaic Virus (T=1) Particle
PDB: 1YC6, www.pdb.org

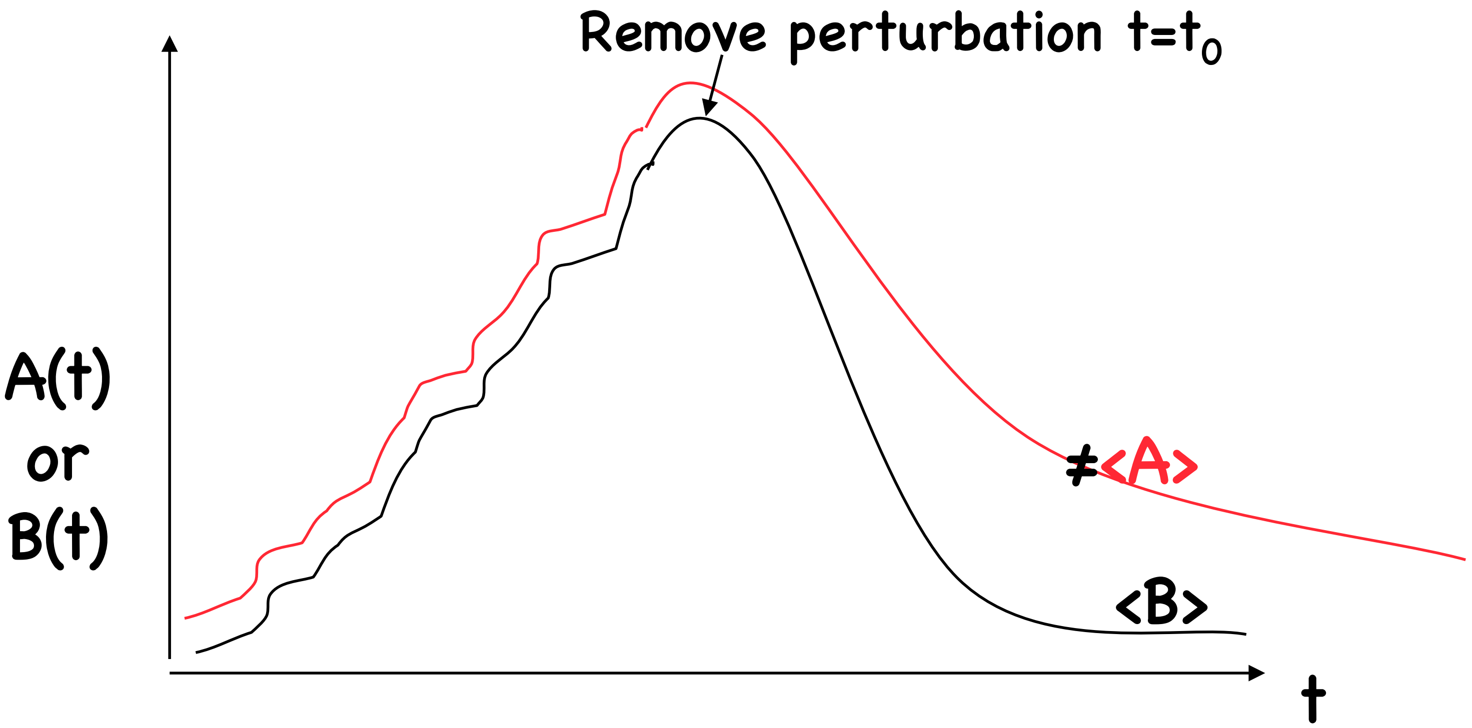
Lecture 1

Separation of Scales in Time and Space

Certain degrees of freedom (position, momentum, etc) of different components of a system (protein, water) come to equilibrium more quickly than do others.

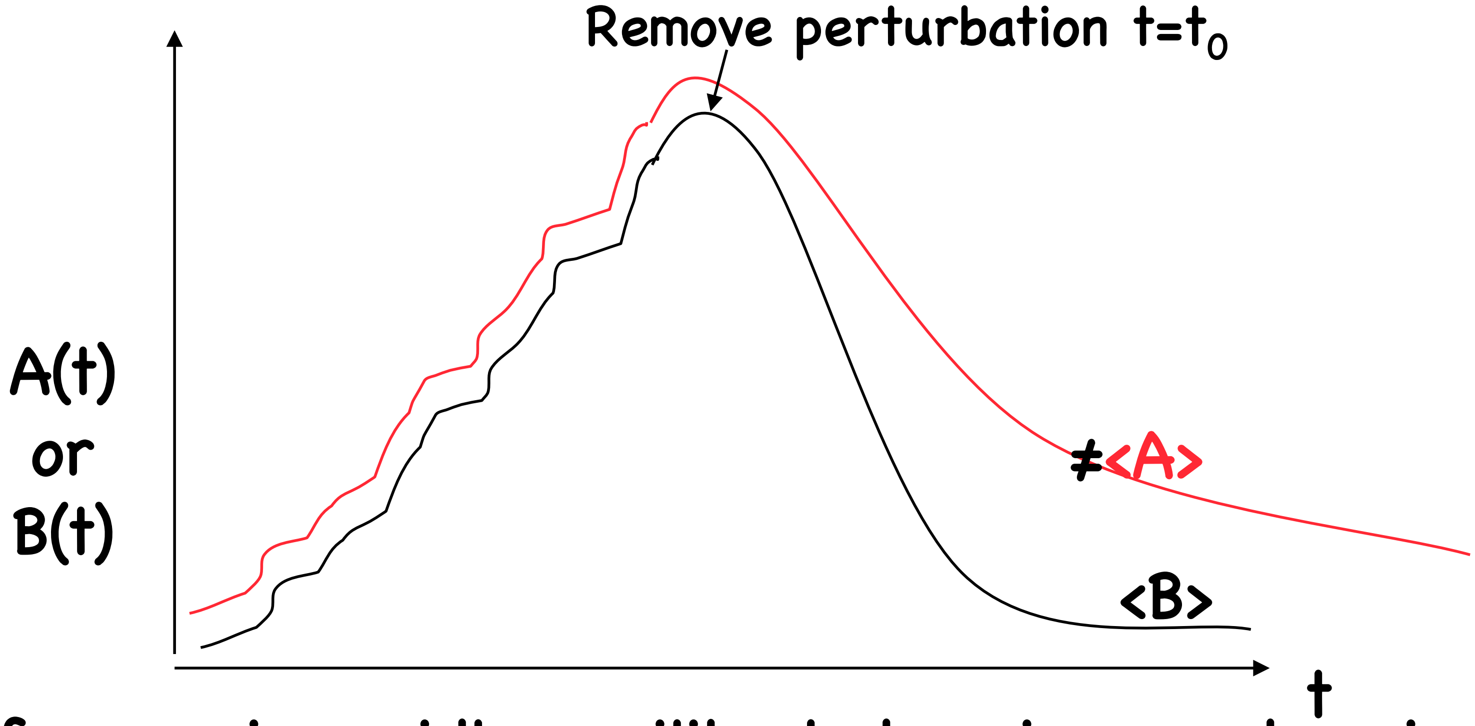
Separation of Scales in Time and Space

Certain degrees of freedom (position, momentum, etc) of different components of a system (protein, water) come to equilibrium more quickly than do others.



Separation of Scales in Time and Space

Certain degrees of freedom (position, momentum, etc) of different components of a system (protein, water) come to equilibrium more quickly than do others.



This difference in rapidly equilibrated regions, and regions whose equilibrium properties are more slowly established, allows for a physically correct separation of spatial scales and time scales.

Coarse-Graining

Given protein (A) and solvent (B) degrees of freedom, lets follow equilibrium property, O, of protein:

$$\langle O \rangle = \frac{\int_{-\infty}^{\infty} d\vec{r}_A^{\,N} \int_{-\infty}^{\infty} d\vec{r}_B^{\,N}\, O\left(\vec{r}_A^{\,N}, \vec{r}_B^{\,N}\right) exp\left[-\beta H\left(\vec{r}_A^{\,N}, \vec{r}_B^{\,N}\right)\right]}{(Q =) \int_{-\infty}^{\infty} d\vec{r}_A^{\,N} \int_{-\infty}^{\infty} d\vec{r}_B^{\,N}\, exp\left[-\beta H\left(\vec{r}_A^{\,N}, \vec{r}_B^{\,N}\right)\right]}$$

Coarse-Graining

Given protein (A) and solvent (B) degrees of freedom, lets follow equilibrium property, O, of protein:

$$\langle O \rangle = \frac{\int_{-\infty}^{\infty} d\vec{r}_A^{\,N} \int_{-\infty}^{\infty} d\vec{r}_B^{\,N}\, O\left(\vec{r}_A^{\,N}, \vec{r}_B^{\,N}\right) exp\left[-\beta H\left(\vec{r}_A^{\,N}, \vec{r}_B^{\,N}\right)\right]}{(Q =) \int_{-\infty}^{\infty} d\vec{r}_A^{\,N} \int_{-\infty}^{\infty} d\vec{r}_B^{\,N}\, exp\left[-\beta H\left(\vec{r}_A^{\,N}, \vec{r}_B^{\,N}\right)\right]}$$

Now lets separate protein (A) and solvent (B) degrees of freedom, for the equilibrium property, O, of protein:

$$Q = \int_{-\infty}^{\infty} d\vec{r}_A^{\,N}\, exp\left[-\beta H_{AA}\left(\vec{r}_A^{\,N}\right)\right]\Big] \int_{-\infty}^{\infty} d\vec{r}_B^{\,N}\, exp\left[-\beta\left(H_{AB}\left(\vec{r}_A^{\,N}, \vec{r}_B^{\,N}\right) + H_{BB}\left(\vec{r}_B^{\,N}\right)\right)\right]$$

Coarse-Graining

Given protein (A) and solvent (B) degrees of freedom, lets follow equilibrium property, O, of protein:

$$\langle O \rangle = \frac{\int_{-\infty}^{\infty} d\vec{r}_A^{\,N} \int_{-\infty}^{\infty} d\vec{r}_B^{\,N} O\left(\vec{r}_A^{\,N}, \vec{r}_B^{\,N}\right) exp\left[-\beta H\left(\vec{r}_A^{\,N}, \vec{r}_B^{\,N}\right)\right]}{(Q =) \int_{-\infty}^{\infty} d\vec{r}_A^{\,N} \int_{-\infty}^{\infty} d\vec{r}_B^{\,N}\, exp\left[-\beta H\left(\vec{r}_A^{\,N}, \vec{r}_B^{\,N}\right)\right]}$$

Now lets separate protein (A) and solvent (B) degrees of freedom, for the equilibrium property, O, of protein:

$$Q = \int_{-\infty}^{\infty} d\vec{r}_A^{\,N}\, exp\left[-\beta H_{AA}\left(\vec{r}_A^{\,N}\right)\right] \int_{-\infty}^{\infty} d\vec{r}_B^{\,N}\, exp\left[-\beta\left(H_{AB}\left(\vec{r}_A^{\,N}, \vec{r}_B^{\,N}\right) + H_{BB}\left(\vec{r}_B^{\,N}\right)\right)\right]$$

And rewrite; Q_B is determined by configuration of protein A

$$Q = \int_{-\infty}^{\infty} d\vec{r}_A^{\,N}\, exp\left[-\beta H_{AA}\left(\vec{r}_A^{\,N}\right)\right] Q_B\left(\vec{r}_A^{\,N}\right)$$

Coarse-Graining

Given protein (A) and solvent (B) degrees of freedom, lets follow equilibrium property, O, of protein:

$$\langle O\rangle = \frac{\int_{-\infty}^{\infty} d\vec{r}_A^{\,N} \int_{-\infty}^{\infty} d\vec{r}_B^{\,N}\, O\left(\vec{r}_A^{\,N},\vec{r}_B^{\,N}\right) exp\left[-\beta H\left(\vec{r}_A^{\,N},\vec{r}_B^{\,N}\right)\right]}{(Q=)\int_{-\infty}^{\infty} d\vec{r}_A^{\,N} \int_{-\infty}^{\infty} d\vec{r}_B^{\,N}\, exp\left[-\beta H\left(\vec{r}_A^{\,N},\vec{r}_B^{\,N}\right)\right]}$$

Now lets separate protein (A) and solvent (B) degrees of freedom, for the equilibrium property, O, of protein:

$$Q = \int_{-\infty}^{\infty} d\vec{r}_A^{\,N}\, exp\left[-\beta H_{AA}\left(\vec{r}_A^{\,N}\right)\right] \int_{-\infty}^{\infty} d\vec{r}_B^{\,N}\, exp\left[-\beta\left(H_{AB}\left(\vec{r}_A^{\,N},\vec{r}_B^{\,N}\right)+H_{BB}\left(\vec{r}_B^{\,N}\right)\right)\right]$$

And rewrite; Q_B is determined by configuration of protein A

$$Q = \int_{-\infty}^{\infty} d\vec{r}_A^{\,N}\, exp\left[-\beta H_{AA}\left(\vec{r}_A^{\,N}\right)\right] Q_B\left(\vec{r}_A^{\,N}\right)$$

Our effective Hamiltonian is now in terms of an equilibrium free energy

$$H_{eff} = H_{AA} - kT\, ln\, Q_B\left(\vec{r}_A^{\,N};N,V,T\right)$$

Coarse-Graining

Given protein (A) and solvent (B) degrees of freedom, lets follow equilibrium property, O, of protein:

$$\langle O\rangle = \frac{\int_{-\infty}^{\infty} d\vec{r}_A^{\,N} \int_{-\infty}^{\infty} d\vec{r}_B^{\,N}\, O\left(\vec{r}_A^{\,N},\vec{r}_B^{\,N}\right) exp\left[-\beta H\left(\vec{r}_A^{\,N},\vec{r}_B^{\,N}\right)\right]}{(Q=)\int_{-\infty}^{\infty} d\vec{r}_A^{\,N} \int_{-\infty}^{\infty} d\vec{r}_B^{\,N}\, exp\left[-\beta H\left(\vec{r}_A^{\,N},\vec{r}_B^{\,N}\right)\right]}$$

Now lets separate protein (A) and solvent (B) degrees of freedom, for the equilibrium property, O, of protein:

$$Q=\int_{-\infty}^{\infty} d\vec{r}_A^{\,N}\, exp\left[-\beta H_{AA}\left(\vec{r}_A^{\,N}\right)\right]\int_{-\infty}^{\infty} d\vec{r}_B^{\,N}\, exp\left[-\beta\left(H_{AB}\left(\vec{r}_A^{\,N},\vec{r}_B^{\,N}\right)+H_{BB}\left(\vec{r}_B^{\,N}\right)\right)\right]$$

And rewrite; Q_B is determined by configuration of protein A

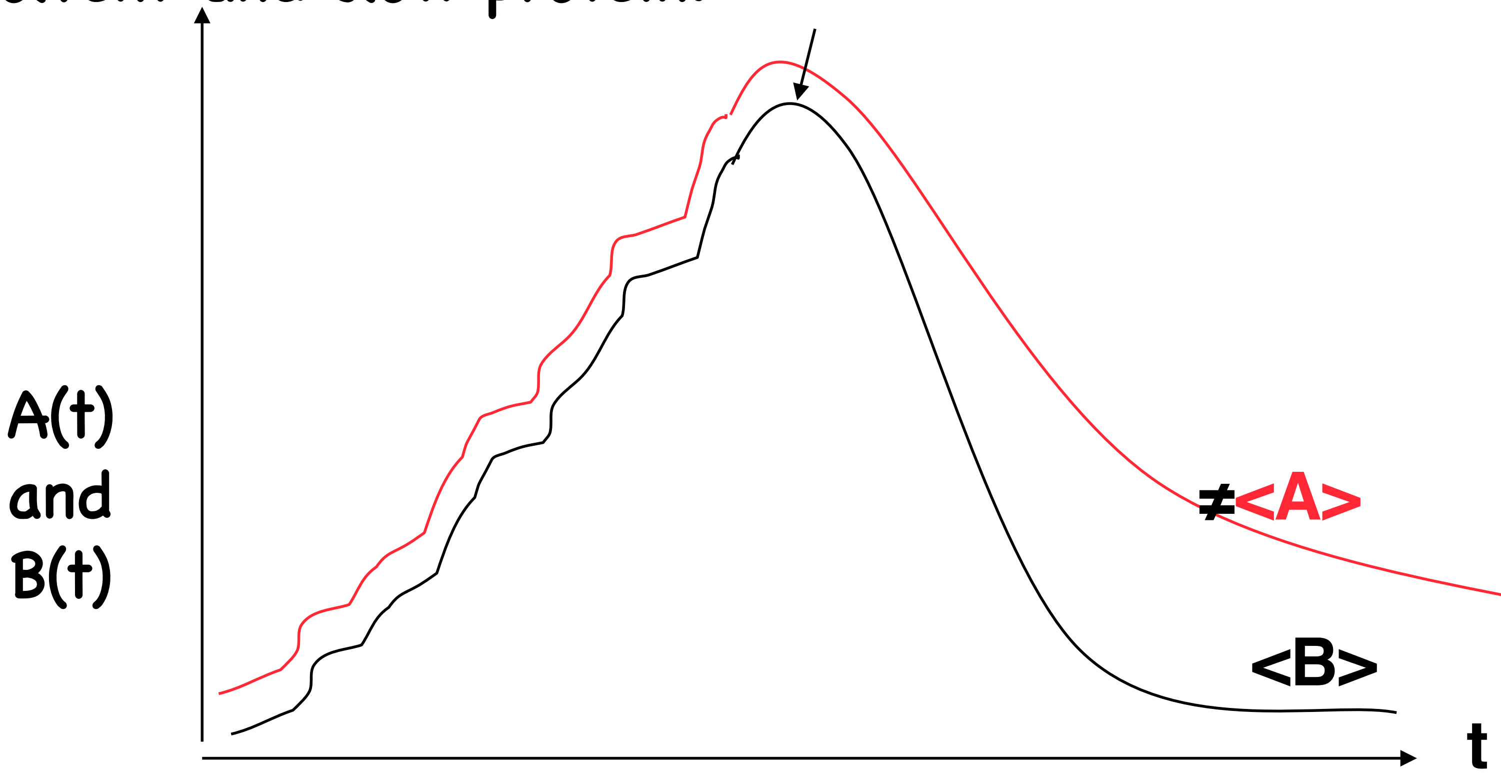
$$Q=\int_{-\infty}^{\infty} d\vec{r}_A^{\,N}\, exp\left[-\beta H_{AA}\left(\vec{r}_A^{\,N}\right)\right]Q_B\left(\vec{r}_A^{\,N}\right)$$

I.e. degrees of freedom belonging to A have not come to equilibrium but B degrees of freedom have!

$$H_{eff}=H_{AA}-kT\,ln\,Q_B\left(\vec{r}_A^{\,N};N,V,T\right)$$

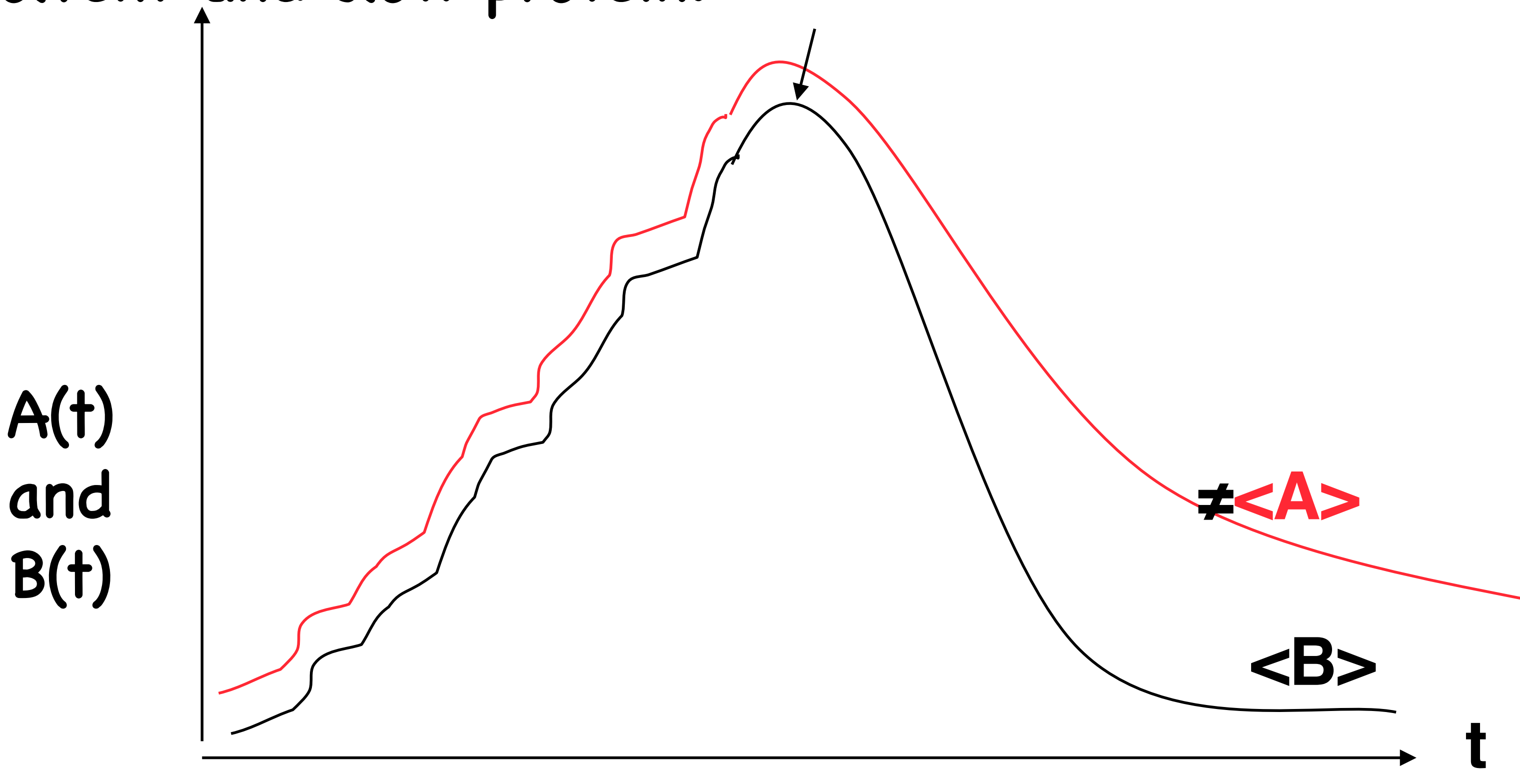
Separation of Timescales

The translational motion of a large protein (A(t)) in a bath of small solvent molecules (B(t)) involves a separation of timescales: fast solvent and slow protein.



Separation of Timescales

The translational motion of a large protein (A(t)) in a bath of small solvent molecules (B(t)) involves a separation of timescales: fast solvent and slow protein.



However our numerical MD must follow the fastest timescale, which is not only costly due to its small Δt, but which is also the least interesting.

Separation of Timescales

Lets ignore the solvent altogether by "integrating it out", but include it implicitly as it must influence the protein's equilibrium relaxation process.

Separation of Timescales

Lets ignore the solvent altogether by “integrating it out”, but include it implicitly as it must influence the protein's equilibrium relaxation process.

Lets start by replacing it with an effective frictional force on the protein.This eliminates the uninteresting fast timescale, that allows us to define a new equation of motion:

$$m\frac{dv}{dt} = -\gamma v$$

Separation of Timescales

Lets ignore the solvent altogether by “integrating it out”, but include it implicitly as it must influence the protein's equilibrium relaxation process.

Lets start by replacing it with an effective frictional force on the protein.This eliminates the uninteresting fast timescale, that allows us to define a new equation of motion:

$$m\frac{dv}{dt} = -\gamma v$$

If we integrate through time, then the solution is

$$v(t) = v(0)exp(-\gamma t / m)$$

Separation of Timescales

Lets ignore the solvent altogether by “integrating it out”, but include it implicitly as it must influence the protein's equilibrium relaxation process.

Lets start by replacing it with an effective frictional force on the protein. This eliminates the uninteresting fast timescale, that allows us to define a new equation of motion:

$$m\frac{dv}{dt} = -\gamma v$$

If we integrate through time, then the solution is

$$v(t) = v(0)\,exp(-\gamma t/m)$$

Note that the limiting velocity goes to zero at $t \to \infty$

Separation of Timescales

However, the protein will eventually relax back to equilibrium, and the limiting velocity fluctuations should be governed by the usual Maxwell-Boltzmann (equipartition of energy),

$$\left\langle v^2 \right\rangle = kT / m$$

Separation of Timescales

However, the protein will eventually relax back to equilibrium, and the limiting velocity fluctuations should be governed by the usual Maxwell-Boltzmann (equipartition of energy),

$$\left\langle v^2 \right\rangle = kT / m$$

Since we get no equipartition of energy (equation on previous slide shows velocities going to zero), this damped motion model is incomplete!

Separation of Timescales

However, the protein will eventually relax back to equilibrium, and the limiting velocity fluctuations should be governed by the usual Maxwell-Boltzmann (equipartition of energy),

$$\langle v^2 \rangle = kT / m$$

Since we get no equipartition of energy (equation on previous slide shows velocities going to zero), this damped motion model is incomplete!

Apparently we are missing an additional force due to solvent. Since solvent is at equilibrium, its motion is only equilibrium fluctuations about its mean.

Fluctuation-Dissipation Theorem

This manifests itself on the protein motion as a random fluctuation force

$$m\frac{dv}{dt} = -\gamma v + \delta Q(t)$$

Fluctuation-Dissipation Theorem

This manifests itself on the protein motion as a random fluctuation force

$$m\frac{dv}{dt} = -\gamma v + \delta Q(t)$$

These random fluctuations at equilibrium, are well-described by a Gaussian distribution

$$\langle \delta Q(t) \rangle = 0 \qquad \langle \delta Q(t)\delta Q(t') \rangle = 2B\delta t$$

Fluctuation-Dissipation Theorem

This manifests itself on the protein motion as a random fluctuation force

$$m\frac{dv}{dt} = -\gamma v + \delta Q(t)$$

These random fluctuations at equilibrium, are well-described by a Gaussian distribution

$$\langle \delta Q(t) \rangle = 0 \qquad \langle \delta Q(t)\delta Q(t') \rangle = 2B\delta t$$

Now what do EOM look like?

Fluctuation-Dissipation Theorem

This manifests itself on the protein motion as a random fluctuation force

$$m\frac{dv}{dt} = -\gamma v + \delta Q(t)$$

These random fluctuations at equilibrium, are well-described by a Gaussian distribution

$$\langle \delta Q(t) \rangle = 0 \qquad \langle \delta Q(t)\delta Q(t') \rangle = 2B\delta t$$

Now what do EOM look like?

Solve ordinary first order differential equation by multiplying with exponential integrating factor:

$$exp(-\gamma t / m)\frac{dv}{dt} - exp(-\gamma t / m)\frac{\gamma v}{m} = \frac{\delta Q(t)}{m} exp(-\gamma t / m)$$

Fluctuation-Dissipation Theorem

$$exp\left(-\gamma t/m\right)\frac{dv}{dt} - exp\left(-\gamma t/m\right)\frac{\gamma v}{m} = \frac{\delta Q\left(t\right)}{m} exp\left(-\gamma t/m\right)$$

Fluctuation-Dissipation Theorem

$$exp(-\gamma t/m)\frac{dv}{dt} - exp(-\gamma t/m)\frac{\gamma v}{m} = \frac{\delta Q(t)}{m} exp(-\gamma t/m)$$

$$\frac{d\left[v(t)\,exp(-\gamma t/m)\right]}{dt} = \frac{\delta Q(t)}{m} exp(-\gamma t/m)$$

Fluctuation-Dissipation Theorem

$$exp\left(-\gamma t / m\right)\frac{dv}{dt} - exp\left(-\gamma t / m\right)\frac{\gamma v}{m} = \frac{\delta Q\left(t\right)}{m}exp\left(-\gamma t / m\right)$$

$$\frac{d\left[v(t)exp\left(-\gamma t / m\right)\right]}{dt} = \frac{\delta Q\left(t\right)}{m}exp\left(-\gamma t / m\right)$$

$$v\left(t\right) = v\left(0\right)exp\left(\gamma t / m\right) + \int_{0}^{t} e^{-\gamma\left(t'-t\right)/m}\delta Q\left(t'\right)dt'$$

Fluctuation-Dissipation Theorem

$$exp\left(-\gamma t/m\right)\frac{dv}{dt} - exp\left(-\gamma t/m\right)\frac{\gamma v}{m} = \frac{\delta Q(t)}{m}exp\left(-\gamma t/m\right)$$

$$\frac{d\left[v(t)exp\left(-\gamma t/m\right)\right]}{dt} = \frac{\delta Q(t)}{m}exp\left(-\gamma t/m\right)$$

$$v(t) = v(0)exp\left(\gamma t/m\right) + \int_0^t e^{-\gamma(t'-t)/m}\delta Q(t')dt'$$

Damping that slows system down

Fluctuation-Dissipation Theorem

$$exp\left(-\gamma t / m\right)\frac{dv}{dt} - exp\left(-\gamma t / m\right)\frac{\gamma v}{m} = \frac{\delta Q\left(t\right)}{m}exp\left(-\gamma t / m\right)$$

$$\frac{d\left[v(t)exp\left(-\gamma t / m\right)\right]}{dt} = \frac{\delta Q\left(t\right)}{m}exp\left(-\gamma t / m\right)$$

$$v\left(t\right) = v\left(0\right)exp\left(\gamma t / m\right) + \int_0^t e^{-\gamma\left(t'-t\right)/m}\delta Q\left(t'\right)dt'$$

Damping that slows system down is now balanced by a random term that speeds the system up

Fluctuation-Dissipation Theorem

$$exp\left(-\gamma t/m\right)\frac{dv}{dt}-exp\left(-\gamma t/m\right)\frac{\gamma v}{m}=\frac{\delta Q\left(t\right)}{m}exp\left(-\gamma t/m\right)$$

$$\frac{d\left[v(t)\,exp\left(-\gamma t/m\right)\right]}{dt}=\frac{\delta Q\left(t\right)}{m}exp\left(-\gamma t/m\right)$$

$$v\left(t\right)=v\left(0\right)exp\left(\gamma t/m\right)+\int_{0}^{t}e^{-\gamma\left(t'-t\right)/m}\,\delta Q\left(t'\right)dt'$$

Damping that slows system down is now balanced by a random term that speeds the system up

The relationship between friction and random force can be determined in limit as t→∞ for which the fluctuations of the velocity should conform to the equipartition of energy

Fluctuation-Dissipation Theorem

Continue solving, and many tedious steps later we arrive at:

$$\left\langle v^2 \right\rangle = v(0)^2 \, exp(-2\gamma t / m) + \frac{B}{\gamma m}\left[1 - exp(-2\gamma t / m)\right]$$

Equipartition of energy allows us to define B

Fluctuation-Dissipation Theorem

Continue solving, and many tedious steps later we arrive at:

$$\langle v^2 \rangle = v(0)^2 \, exp(-2\gamma t / m) + \frac{B}{\gamma m}\left[1 - exp(-2\gamma t / m)\right]$$

Equipartition of energy allows us to define B

$$\langle v^2 \rangle = kT / m \qquad \rightarrow \qquad B = \gamma kT$$

Fluctuation-Dissipation Theorem

Continue solving, and many tedious steps later we arrive at:

$$\langle v^2 \rangle = v(0)^2 \exp(-2\gamma t / m) + \frac{B}{\gamma m}\left[1 - \exp(-2\gamma t / m)\right]$$

Equipartition of energy allows us to define B

$$\langle v^2 \rangle = kT / m \qquad \rightarrow \qquad B = \gamma kT$$

And 2nd moment of Gaussian random force

$$\langle \delta Q(t) \delta Q(t') \rangle = 2\gamma kT \delta t$$

Fluctuation-Dissipation Theorem

Continue solving, and many tedious steps later we arrive at:

$$\left\langle v^2 \right\rangle = v(0)^2 \exp(-2\gamma t / m) + \frac{B}{\gamma m}\left[1 - \exp(-2\gamma t / m)\right]$$

Equipartition of energy allows us to define B

$$\left\langle v^2 \right\rangle = kT / m \qquad \rightarrow \qquad B = \gamma kT$$

And 2nd moment of Gaussian random force

$$\left\langle \delta Q(t) \delta Q(t') \right\rangle = 2\gamma kT \delta t$$

The essence of fluctuation-dissipation theorem is that it relates equilibrium thermal fluctuations to out-of-equilibrium quantities.

Fluctuation-Dissipation Theorem

Continue solving, and many tedious steps later we arrive at:

$$\left\langle v^2 \right\rangle = v(0)^2 \, exp(-2\gamma t / m) + \frac{B}{\gamma m}\left[1 - exp(-2\gamma t / m)\right]$$

Equipartition of energy allows us to define B

$$\left\langle v^2 \right\rangle = kT / m \qquad \rightarrow \qquad B = \gamma kT$$

And 2nd moment of Gaussian random force

$$\left\langle \delta Q(t) \delta Q(t') \right\rangle = 2\gamma kT \delta t$$

The essence of fluctuation-dissipation theorem is that it relates equilibrium thermal fluctuations to out-of-equilibrium quantities.

It states that the dissipative effects needed to restore a non-equilibrium system to equilibrium are derived from equilibrium fluctuations of our system correlated at different times.

Newtonian to Langevin Dynamics

Newton's equation of motion of all degrees of freedom:

$$\vec{f}(t) = m\frac{d\vec{v}(t)}{dt}$$

Newtonian to Langevin Dynamics

Newton's equation of motion of all degrees of freedom:

$$m\frac{d\vec{v}(t)}{dt} = \vec{f}(t)$$

Replaced with Langevin equation for slow particle, with fast degrees of freedom manifested as effective forces:

$$m\frac{d\vec{v}(t)}{dt} = \vec{f}(t) - \frac{1}{D^*}\vec{v}(t) + \vec{Q}(t)$$

Newtonian to Langevin Dynamics

Newton's equation of motion of all degrees of freedom:

$$m\frac{d\vec{v}(t)}{dt} = \vec{f}(t)$$

Replaced with Langevin equation for slow particle, with fast degrees of freedom manifested as effective forces:

$$m\frac{d\vec{v}(t)}{dt} = \vec{f}(t) - \frac{1}{D*}\vec{v}(t) + \vec{Q}(t)$$

(1) A friction or drag force representing hydrodynamic effects of the missing solvent. D* is the mobility; D*=1/(6πηa), where a is particle radius

$$-\frac{1}{D*}\vec{v}(t)$$

Newtonian to Langevin Dynamics

Newton's equation of motion of all degrees of freedom:

$$m\frac{d\vec{v}(t)}{dt} = \vec{f}(t)$$

Replaced with Langevin equation for slow particle, with fast degrees of freedom manifested as effective forces:

$$m\frac{d\vec{v}(t)}{dt} = \vec{f}(t) - \frac{1}{D*}\vec{v}(t) + \vec{Q}(t)$$

(1) A friction or drag force representing hydrodynamic effects of the missing solvent. D* is the mobility; D*=1/(6πηa), where a is particle radius

$$-\frac{1}{D*}\vec{v}(t)$$

(2) Random fluctuating force described as "white noise" or Gaussian distribution of the random walk

$$\vec{Q}(t) \quad \rightarrow \quad \langle \vec{Q} \rangle = 0 \qquad \langle \vec{Q}^2 \rangle \propto D$$

Langevin Dynamics

Let's examine this separation of into fast and slow timescales

$$m\dot{\vec{v}} = -\frac{1}{D^*}\vec{v} + \vec{Q}$$

Ignore systematic force.

Langevin Dynamics

Let's examine this separation of into fast and slow timescales

$$m\dot{\vec{v}} = -\frac{1}{D^*}\vec{v} + \vec{Q}$$

Ignore systematic force.

$$m\ddot{\vec{r}} = -\frac{1}{D^*}\dot{\vec{r}} + \vec{Q}$$

Rewrite in terms of position variable

Langevin Dynamics

Let's examine this separation of into fast and slow timescales

$$m\dot{\vec{v}} = -\frac{1}{D*}\vec{v} + \vec{Q}$$

Ignore systematic force.

$$m\ddot{\vec{r}} = -\frac{1}{D*}\dot{\vec{r}} + \vec{Q}$$

Rewrite in terms of position variable

$$m\vec{r}\ddot{\vec{r}} = -\frac{1}{D*}\vec{r}\dot{\vec{r}} + \vec{r}\vec{Q}$$

Multiply by position variable r. Eq. (1)

Langevin Dynamics

Let's examine this separation of into fast and slow timescales

$$m\dot{\vec{v}} = -\frac{1}{D^*}\vec{v} + \vec{Q}$$

Ignore systematic force.

$$m\ddot{\vec{r}} = -\frac{1}{D^*}\dot{\vec{r}} + \vec{Q}$$

Rewrite in terms of position variable

$$m\vec{r}\ddot{\vec{r}} = -\frac{1}{D^*}\vec{r}\dot{\vec{r}} + \vec{r}\vec{Q}$$

Multiply by position variable r. Eq. (1)

$$\vec{r}\ddot{\vec{r}} = \frac{1}{2}\frac{d^2\left(\vec{r}^{\,2}\right)}{dt^2} - \dot{\vec{r}}^{\,2} \qquad \vec{r}\dot{\vec{r}} = \frac{1}{2}\frac{d\left(\vec{r}^{\,2}\right)}{dt}$$

Definitions

Langevin Dynamics

Let's examine this separation of into fast and slow timescales

$$m\dot{\vec{v}} = -\frac{1}{D^*}\vec{v} + \vec{Q}$$

Ignore systematic force.

$$m\ddot{\vec{r}} = -\frac{1}{D^*}\dot{\vec{r}} + \vec{Q}$$

Rewrite in terms of position variable

$$m\vec{r}\ddot{\vec{r}} = -\frac{1}{D^*}\vec{r}\dot{\vec{r}} + \vec{r}\vec{Q}$$

Multiply by position variable r. Eq. (1)

$$\vec{r}\ddot{\vec{r}} = \frac{1}{2}\frac{d^2\left(\vec{r}^{\,2}\right)}{dt^2} - \dot{\vec{r}}^{\,2} \qquad \vec{r}\dot{\vec{r}} = \frac{1}{2}\frac{d\left(\vec{r}^{\,2}\right)}{dt}$$

Definitions

Substitute definitions in Eq. (1) and take thermal average:

Langevin Dynamics

Let's examine this separation of into fast and slow timescales

$$m\dot{\vec{v}} = -\frac{1}{D*}\vec{v} + \vec{Q}$$

Ignore systematic force.

$$m\ddot{\vec{r}} = -\frac{1}{D*}\dot{\vec{r}} + \vec{Q}$$

Rewrite in terms of position variable

$$m\vec{r}\ddot{\vec{r}} = -\frac{1}{D*}\vec{r}\dot{\vec{r}} + \vec{r}\vec{Q}$$

Multiply by position variable r. Eq. (1)

$$\vec{r}\ddot{\vec{r}} = \frac{1}{2}\frac{d^2\left(\vec{r}^{\,2}\right)}{dt^2} - \dot{\vec{r}}^{\,2} \qquad \vec{r}\dot{\vec{r}} = \frac{1}{2}\frac{d\left(\vec{r}^{\,2}\right)}{dt}$$

Definitions

Substitute definitions in Eq. (1) and take thermal average:

$$m\frac{d^2\langle\vec{r}\cdot\vec{r}\rangle}{dt^2} - 2m\langle\dot{\vec{r}}\cdot\dot{\vec{r}}\rangle = -\frac{1}{D*}\frac{d\langle\vec{r}\cdot\vec{r}\rangle}{dt} + 2\langle\vec{r}\cdot\vec{Q}\rangle$$

Langevin Dynamics

$$m\frac{d^2\langle\vec{r}\cdot\vec{r}\rangle}{dt^2}+\frac{1}{D^*}\frac{d\langle\vec{r}\cdot\vec{r}\rangle}{dt}=2m\langle\dot{\vec{r}}\cdot\dot{\vec{r}}\rangle+2\langle\vec{r}\cdot\vec{Q}\rangle$$

Assume that slow particle is at thermal equilibrium with solvent bath so equipartition holds

Langevin Dynamics

$$\frac{d^2\langle \vec{r}\cdot\vec{r}\rangle}{dt^2}+\frac{1}{mD^*}\frac{d\langle \vec{r}\cdot\vec{r}\rangle}{dt}=\cancel{2\langle \dot{\vec{r}}\cdot\dot{\vec{r}}\rangle}+\frac{2}{m}\langle \vec{r}\cdot\vec{Q}\rangle$$

Assume that slow particle is at thermal equilibrium with solvent bath so equipartition holds

$$\frac{d^2\langle \vec{r}\cdot\vec{r}\rangle}{dt^2}+\frac{1}{mD^*}\frac{d\langle \vec{r}\cdot\vec{r}\rangle}{dt}=\frac{6kT}{m}+\frac{2}{m}\langle \vec{r}\cdot\vec{Q}\rangle$$

Langevin Dynamics

$$\frac{d^2\langle \vec{r}\cdot\vec{r}\rangle}{dt^2}+\frac{1}{mD^*}\frac{d\langle \vec{r}\cdot\vec{r}\rangle}{dt}=2\langle \dot{\vec{r}}\cdot\dot{\vec{r}}\rangle+\frac{2}{m}\langle \vec{r}\cdot\vec{Q}\rangle$$

Assume that slow particle is at thermal equilibrium with solvent bath so equipartition holds

$$\frac{d^2\langle \vec{r}\cdot\vec{r}\rangle}{dt^2}+\frac{1}{mD^*}\frac{d\langle \vec{r}\cdot\vec{r}\rangle}{dt}=\frac{6kT}{m}+\frac{2}{m}\langle \vec{r}\cdot\vec{Q}\rangle$$

Position and random force are uncorrelated so <r·Q>=0

$$\frac{d^2\langle \vec{r}\cdot\vec{r}\rangle}{dt^2}+\frac{1}{mD^*}\frac{d\langle \vec{r}\cdot\vec{r}\rangle}{dt}=\frac{6kT}{m}$$

Langevin Dynamics

$$\frac{d^2\langle\vec{r}\cdot\vec{r}\rangle}{dt^2}+\frac{1}{mD^*}\frac{d\langle\vec{r}\cdot\vec{r}\rangle}{dt}=2\langle\dot{\vec{r}}\cdot\dot{\vec{r}}\rangle+\frac{2}{m}\langle\vec{r}\cdot\vec{Q}\rangle$$

Assume that slow particle is at thermal equilibrium with solvent bath so equipartition holds

$$\frac{d^2\langle\vec{r}\cdot\vec{r}\rangle}{dt^2}+\frac{1}{mD^*}\frac{d\langle\vec{r}\cdot\vec{r}\rangle}{dt}=\frac{6kT}{m}+\frac{2}{m}\langle\vec{r}\cdot\vec{Q}\rangle$$

Position and random force are uncorrelated so <r·Q>=0

$$\frac{d^2\langle\vec{r}\cdot\vec{r}\rangle}{dt^2}+\frac{1}{mD^*}\frac{d\langle\vec{r}\cdot\vec{r}\rangle}{dt}=\frac{6kT}{m}$$

And now this partial differential can be solved

$$\langle\vec{r}\cdot\vec{r}\rangle=6kTmD^{*2}\left\{\frac{t}{mD^*}-\left(1-e^{-t/mD^*}\right)\right\}$$

Langevin Dynamics: Limits

Define a time scale as the ratio of mass over friction:

$$\tau = mD*$$

Langevin Dynamics: Limits

Define a time scale as the ratio of mass over friction:

$$\tau = mD^*$$

And substitute

$$\langle \vec{r} \cdot \vec{r} \rangle = \frac{6kT\tau^2}{m}\left\{ \frac{t}{\tau} - \left(1 - e^{-t/\tau}\right) \right\}$$

Langevin Dynamics: Limits

Define a time scale as the ratio of mass over friction:

$$\tau = mD^*$$

And substitute

$$\langle \vec{r} \cdot \vec{r} \rangle = \frac{6kT\tau^2}{m} \left\{ \frac{t}{\tau} - \left(1 - e^{-t/\tau} \right) \right\}$$

For times much shorter than this time variable:

$$t << \tau$$

Langevin Dynamics: Limits

Define a time scale as the ratio of mass over friction:

$$\tau = mD^*$$

And substitute

$$\langle \vec{r} \cdot \vec{r} \rangle = \frac{6kT\tau^2}{m} \left\{ \frac{t}{\tau} - \left(1 - e^{-t/\tau} \right) \right\}$$

For times much shorter than this time variable:

$$t << \tau$$

We recover an MD like expression (we ignored systematic force):

$$\langle \vec{r} \cdot \vec{r} \rangle = \frac{3kTt^2}{m} = m \langle \vec{v} \cdot \vec{v} \rangle t^2 \xrightarrow{\qquad} \vec{r} = \vec{v}t$$

Langevin Dynamics: Limits

Define a time scale as the ratio of mass over friction:

$$\tau = mD^*$$

And substitute

$$\langle \vec{r} \cdot \vec{r} \rangle = \frac{6kT\tau^2}{m}\left\{\frac{t}{\tau} - \left(1 - e^{-t/\tau}\right)\right\}$$

For times much shorter than this time variable:

$$t << \tau$$

We recover an MD like expression (we ignored systematic force):

$$\langle \vec{r} \cdot \vec{r} \rangle = \frac{3kTt^2}{m} = m\langle \vec{v} \cdot \vec{v} \rangle t^2 \longrightarrow \vec{r} = \vec{v}t$$

For times much longer than this time variable:

$$t >> \tau$$

Langevin Dynamics: Limits

Define a time scale as the ratio of mass over friction:

$$\tau = mD^*$$

And substitute

$$\langle \vec{r}\cdot\vec{r}\rangle = \frac{6kT\tau^2}{m}\left\{\frac{t}{\tau} - \left(1 - e^{-t/\tau}\right)\right\}$$

For times much shorter than this time variable:

$$t << \tau$$

We recover an MD like expression (we ignored systematic force):

$$\langle \vec{r}\cdot\vec{r}\rangle = \frac{3kTt^2}{m} = m\langle \vec{v}\cdot\vec{v}\rangle t^2 \xrightarrow{\qquad\qquad} \quad \vec{r} = \vec{v}t$$

For times much longer than this time variable:

$$t >> \tau$$

We recover the diffusion expression and fluctuation-dissipation:

$$\langle \vec{r}\cdot\vec{r}\rangle = \frac{6kTt\tau}{m} = 6kTD^*t = 6Dt \qquad D = ktD^* = \frac{kT}{6\pi\eta a}$$

Brownian Dynamics

Brownian dynamics is a next level of coarse-graining. We don't see the instantaneous acceleration of the Brownian particle: i.e. we are on an effectively longer timescale.

$$0 = -\varepsilon\vec{v}(t) + \vec{f}(t) + \vec{Q}(t)$$

$$\vec{v}(t) = \frac{1}{\varepsilon}\left[\vec{f}(t) + \vec{Q}(t)\right]$$

Brownian Dynamics

Brownian dynamics is a next level of coarse-graining. We don't see the instantaneous acceleration of the Brownian particle: i.e. we are on an effectively longer timescale.

$$0 = -\varepsilon \vec{v}(t) + \vec{f}(t) + \vec{Q}(t)$$

$$\vec{v}(t) = \frac{1}{\varepsilon}\left[\vec{f}(t) + \vec{Q}(t)\right]$$

Use simple finite difference expression:

$$r(t + \Delta t) = r(t) + \frac{Dm}{kT}\vec{f}(t)\Delta t + \vec{Q}'(t) \qquad D = \frac{kT}{6\pi\eta a}$$

Brownian Dynamics

Brownian dynamics is a next level of coarse-graining. We don't see the instantaneous acceleration of the Brownian particle: i.e. we are on an effectively longer timescale.

$$0 = -\varepsilon\vec{v}(t) + \vec{f}(t) + \vec{Q}(t)$$

$$\vec{v}(t) = \frac{1}{\varepsilon}\left[\vec{f}(t) + \vec{Q}(t)\right]$$

Use simple finite difference expression:

$$r(t+\Delta t) = r(t) + \frac{Dm}{kT}\vec{f}(t)\Delta t + \vec{Q}'(t)$$

$$D = \frac{kT}{6\pi\eta a}$$

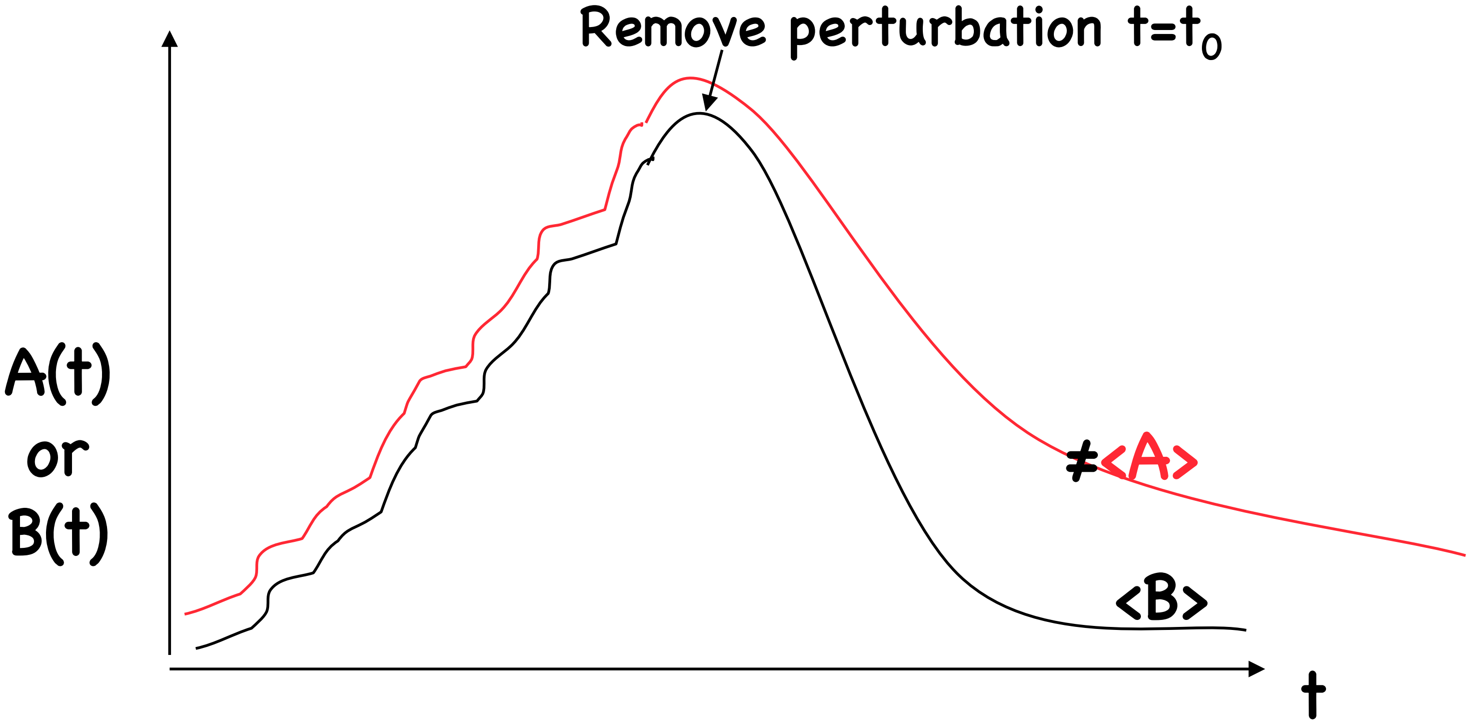
$$\left\langle \vec{Q}'(t) \right\rangle = 0$$

$$\left\langle \vec{Q}'(t)^2 \right\rangle = 6D\Delta t$$

Draw from a Gaussian distribution of this 2nd moment

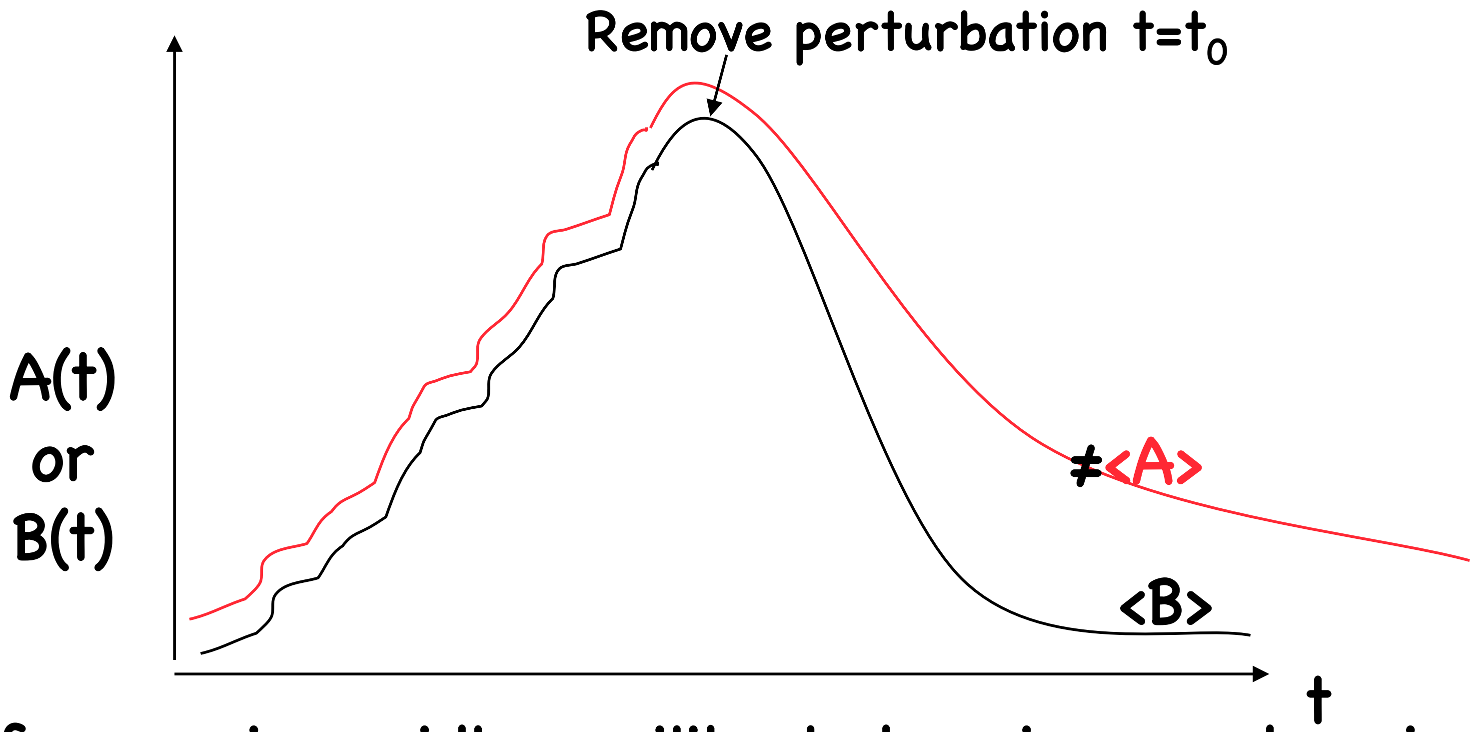
Summary of Coarse Graining

Certain degrees of freedom (position, momentum, etc) of different components of a system (protein, water) come to equilibrium more quickly than do others.



Summary of Coarse Graining

Certain degrees of freedom (position, momentum, etc) of different components of a system (protein, water) come to equilibrium more quickly than do others.



This difference in rapidly equilibrated regions, and regions whose equilibrium properties are more slowly established, allows for a physically correct separation of spatial scales and time scales.

Separation of Lengthscales

How do we decide what to integrate out?

> The work in this paper clearly shows that there is an art in knowing what needs to be preserved in the coarse-graining procedure in order to correctly render the key underlying physical processes one is trying to emulate

M. E. Johnson, T. Head-Gordon, A. A. Louis (2007). Representability problems for coarse-grained water models. J. Chem. Phys. 126, 144509-144519.

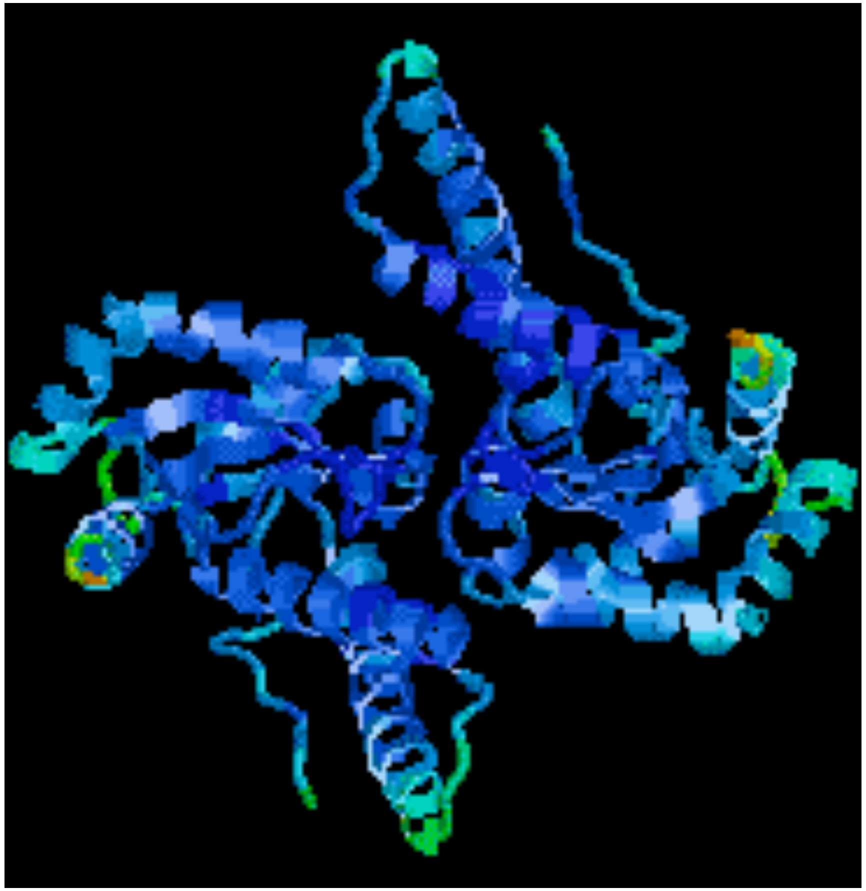
CG electrostatic models for long-range recognition in protein-ligand and protein-protein association

Structural coarse-graining, Henderson uniqueness theorem, and Thermodynamic-dynamic relation for water

Principle of minimal frustration: protein folding, protein aggregation, and protein-membrane interactions

Protein-Ligand Binding: SOD/Superoxide

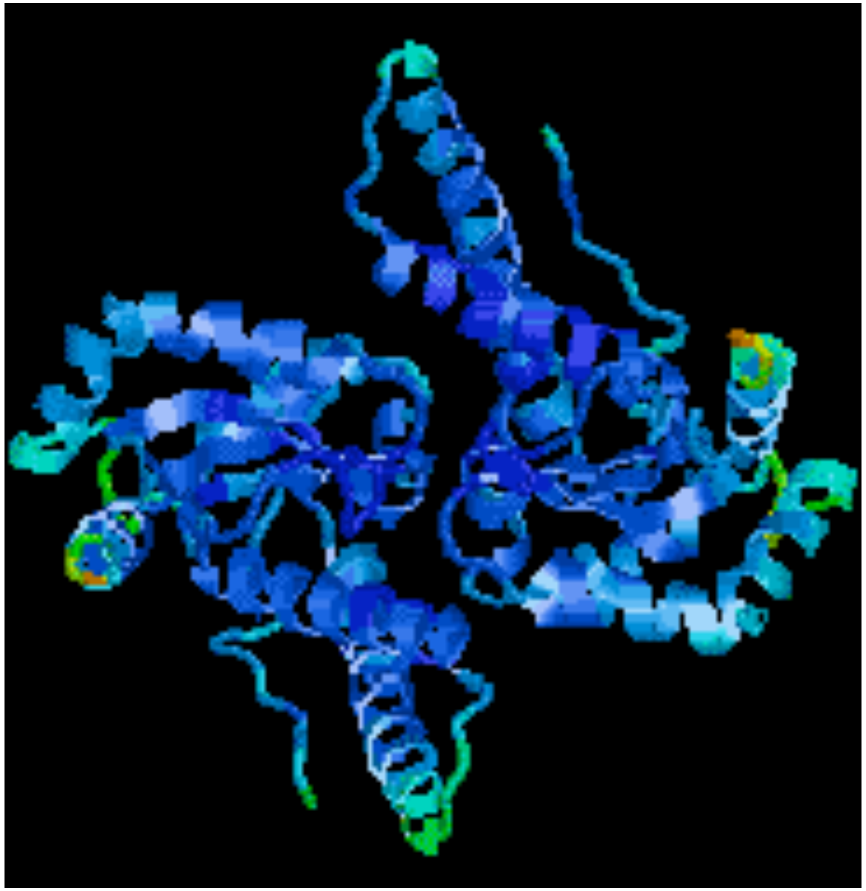
“CuZn superoxide dismutase (CuZnSOD) catalyses the dismutation reaction of the toxic superoxide radical to molecular oxygen and hydrogen peroxide and thus forms a crucial part of the cellular antioxidant defense mechanism”. http://srs.dl.ac.uk/mbg/sod.html



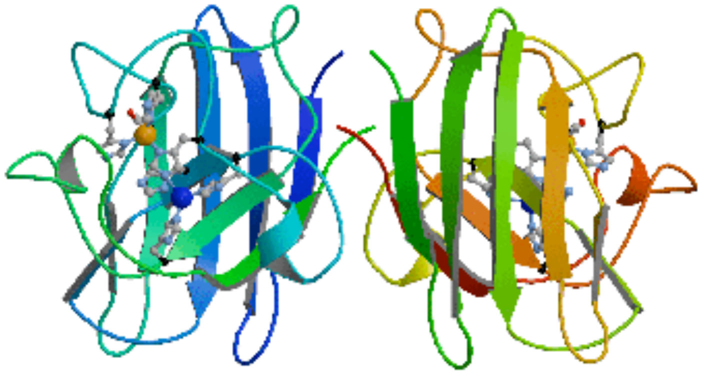
Protein-Ligand Binding: SOD/Superoxide

"CuZn superoxide dismutase (CuZnSOD) catalyses the dismutation reaction of the toxic superoxide radical to molecular oxygen and hydrogen peroxide and thus forms a crucial part of the cellular antioxidant defence mechanism". http://srs.dl.ac.uk/mbg/sod.html

Single point mutations in SOD are associated with the development of a form of motor neuron disease known as ALS or Lou Gehrig's disease



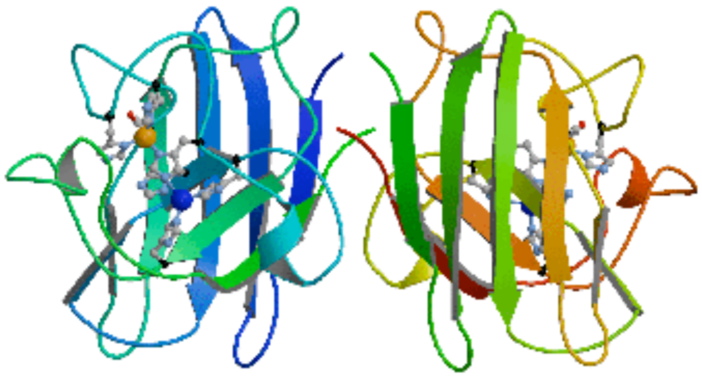
Superoxide Dismutase



CuZnSOD is a protein dimer. Each 151 amino acid monomer folds as an eight stranded greek-key beta-barrel connected by three external loops. The overall charge of SOD is -4e.

http://srs.dl.ac.uk/mbg/sod.html

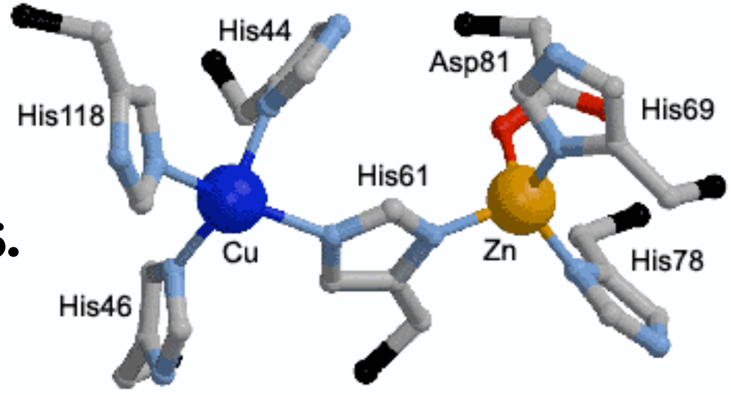
Superoxide Dismutase



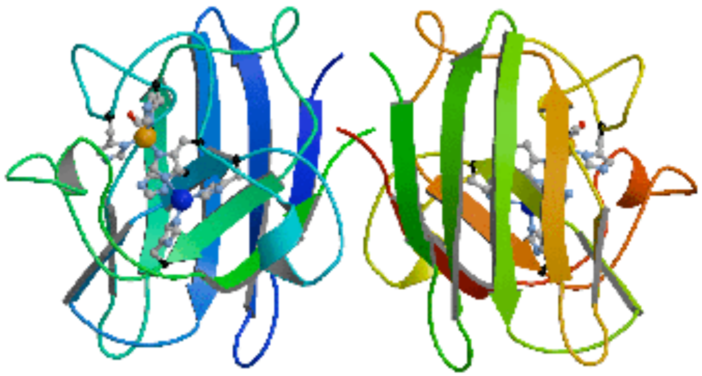
CuZnSOD is a protein dimer. Each 151 amino acid monomer folds as an eight stranded greek-key beta-barrel connected by three external loops. The overall charge of SOD is -4e.

http://srs.dl.ac.uk/mbg/sod.html

The active site of SOD lies at base of 15Å deep cavity formed by two loops containing charged residues. The active site consists of 1 Cu and 1 Zn ion bridged by imidazole ring of His61, a feature unique to enzymes of this class.



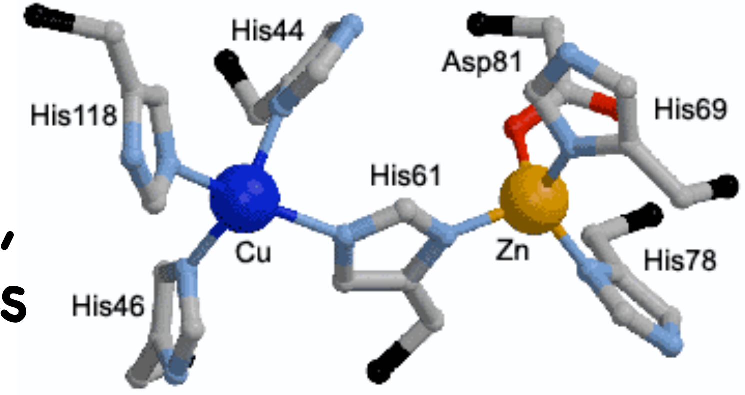
Superoxide Dismutase



CuZnSOD is a protein dimer. Each 151 amino acid monomer folds as an eight stranded greek-key beta-barrel connected by three external loops. The overall charge of SOD is -4e.

http://srs.dl.ac.uk/mbg/sod.html

The active site of SOD lies at base of 15Å deep cavity formed by two loops containing charged residues. The active site Cu is coordinated by His61 and three His ligands and a water molecule, while Zn is ligated by His61 and two His and an aspartic acid.



Diffusion Controlled Chemical Reactions

The overall rate of any (bio)chemical reaction is the rate of the slowest step. Diffusion-controlled reactions are reactions in which the catalysis step occurs so quickly that the reaction rate is determined by how fast reactants encounter each other.

Diffusion Controlled Chemical Reactions

The overall rate of any (bio)chemical reaction is the rate of the slowest step. Diffusion-controlled reactions are reactions in which the catalysis step occurs so quickly that the reaction rate is determined by how fast reactants encounter each other.

Diffusion-controlled rates are more likely found in solution where diffusion of reactants is slower due to the greater number of collisions with solvent molecules, and is typical for evolutionarily optimized enzymes.

Diffusion Controlled Chemical Reactions

The overall rate of any (bio)chemical reaction is the rate of the slowest step. Diffusion-controlled reactions are reactions in which the catalysis step occurs so quickly that the reaction rate is determined by how fast reactants encounter each other.

Diffusion-controlled rates are more likely found in solution where diffusion of reactants is slower due to the greater number of collisions with solvent molecules, and is typical for evolutionarily optimized enzymes.

Experimentally we know the following about the functional properties of the SOD enzyme. The actual reaction itself

Diffusion Controlled Chemical Reactions

The overall rate of any (bio)chemical reaction is the rate of the slowest step. Diffusion-controlled reactions are reactions in which the catalysis step occurs so quickly that the reaction rate is determined by how fast reactants encounter each other.

Diffusion-controlled rates are more likely found in solution where diffusion of reactants is slower due to the greater number of collisions with solvent molecules, and is typical for evolutionarily optimized enzymes.

Experimentally we know the following about the functional properties of the SOD enzyme. The actual reaction itself

$$\text{SOD-Cu}^{2+} + O_2^{\cdot -} \Rightarrow \text{SOD-Cu}^{1+} + O_2$$

$$\text{SOD-Cu}^{1+} + O_2^{\cdot -} + 2H^+ \Rightarrow \text{SOD-Cu}^{2+} + H_2O_2$$

Diffusion Controlled Chemical Reactions

The overall rate of any (bio)chemical reaction is the rate of the slowest step. Diffusion-controlled reactions are reactions in which the catalysis step occurs so quickly that the reaction rate is determined by how fast reactants encounter each other.

Diffusion-controlled rates are more likely found in solution where diffusion of reactants is slower due to the greater number of collisions with solvent molecules, and is typical for evolutionarily optimized enzymes.

Experimentally we know the following about the functional properties of the SOD enzyme. The actual reaction itself

$$\mathrm{SOD\text{-}Cu^{2+}} + \mathrm{O_2^{\cdot -}} \Rightarrow \mathrm{SOD\text{-}Cu^{1+}} + \mathrm{O_2}$$

$$\mathrm{SOD\text{-}Cu^{1+}} + \mathrm{O_2^{\cdot -}} + 2\mathrm{H^+} \Rightarrow \mathrm{SOD\text{-}Cu^{2+}} + \mathrm{H_2O_2}$$

takes very little time to complete, but instead the rate limiting step is diffusion of the O_2^- radical to the active site.

SOD/O_2^- Diffusion Controlled Rates

Thus the SOD catalysis rate of conversion of superoxide to hydrogen peroxide is close to diffusion-controlled. Diffusion controlled rates imply that there is no systematic force between SOD and O_2^- (its all random collisions and friction forces).

This is surprising for several reasons:

SOD/O_2^- Diffusion Controlled Rates

Thus the SOD catalysis rate of conversion of superoxide to hydrogen peroxide is close to diffusion-controlled. Diffusion controlled rates imply that there is no systematic force between SOD and O_2^- (its all random collisions and friction forces).

This is surprising for several reasons:

(1) The overall charge of the protein (-4e) and the O_2^- ligand (-1e) should be a repulsive interaction that would slow down the rate.

SOD/O_2^- Diffusion Controlled Rates

Thus the SOD catalysis rate of conversion of superoxide to hydrogen peroxide is close to diffusion-controlled. Diffusion controlled rates imply that there is no systematic force between SOD and O_2^- (its all random collisions and friction forces).

This is surprising for several reasons:

(1) The overall charge of the protein (-4e) and the O_2^- ligand (-1e) should be a repulsive interaction that would slow down the rate.

(2) The rate decreases with increasing salt concentration

SOD/O_2^- Diffusion Controlled Rates

Thus the SOD catalysis rate of conversion of superoxide to hydrogen peroxide is close to diffusion-controlled. Diffusion controlled rates imply that there is no systematic force between SOD and O_2^- (its all random collisions and friction forces).

This is surprising for several reasons:

(1) The overall charge of the protein (-4e) and the O_2^- ligand (-1e) should be a repulsive interaction that would slow down the rate.

(2) The rate decreases with increasing salt concentration

(3) The active site sits at the bottom of a deep cleft that occupies 1/150th of the total protein surface area.

SOD/O_2^- Diffusion Controlled Rates

Today we will develop a model for the potential energy surface of SOD/O_2^- interaction that includes electrostatics, dielectric boundaries, salt effects (ionic strength of solution), and active site accessibility.

SOD/O_2^- Diffusion Controlled Rates

Today we will develop a model for the potential energy surface of SOD/O_2^- interaction that includes electrostatics, dielectric boundaries, salt effects (ionic strength of solution), and active site accessibility.

We then will use Brownian dynamics to simulate a reaction rate with above systematic forces, combined with a flux model to account for diffusion in an infinite domain. We will compare this to the rate with varying systematic forces (with and without dielectric discontinuities, with and without salt, etc).

SOD/O_2^- Diffusion Controlled Rates

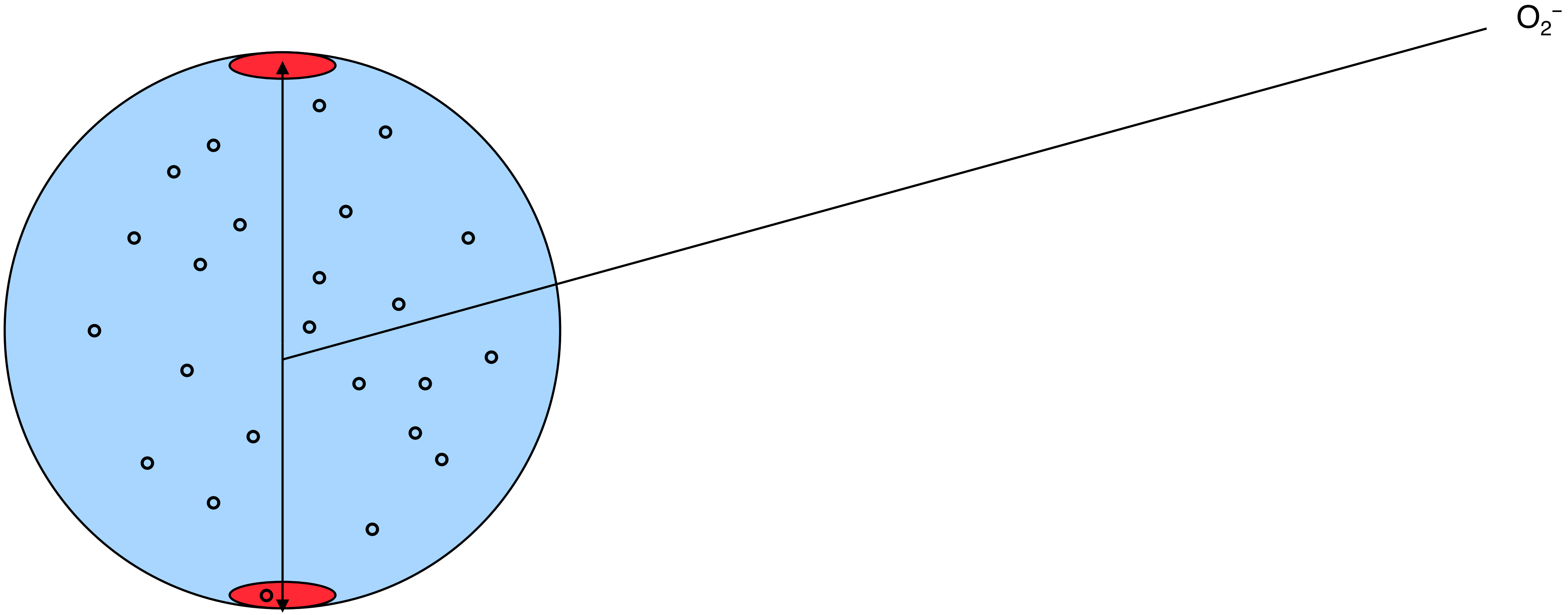
Today we will develop a model for the potential energy surface of SOD/O_2^- interaction that includes electrostatics, dielectric boundaries, salt effects (ionic strength of solution), and active site accessibility.

We then will use Brownian dynamics to simulate a reaction rate with above systematic forces, combined with a flux model to account for diffusion in an infinite domain. We will compare this to the rate with varying systematic forces (with and without dielectric discontinuities, with and without salt, etc).

We propose a possible explanation for poor function for ALS (although recently it has been classified as an aggregation disease)

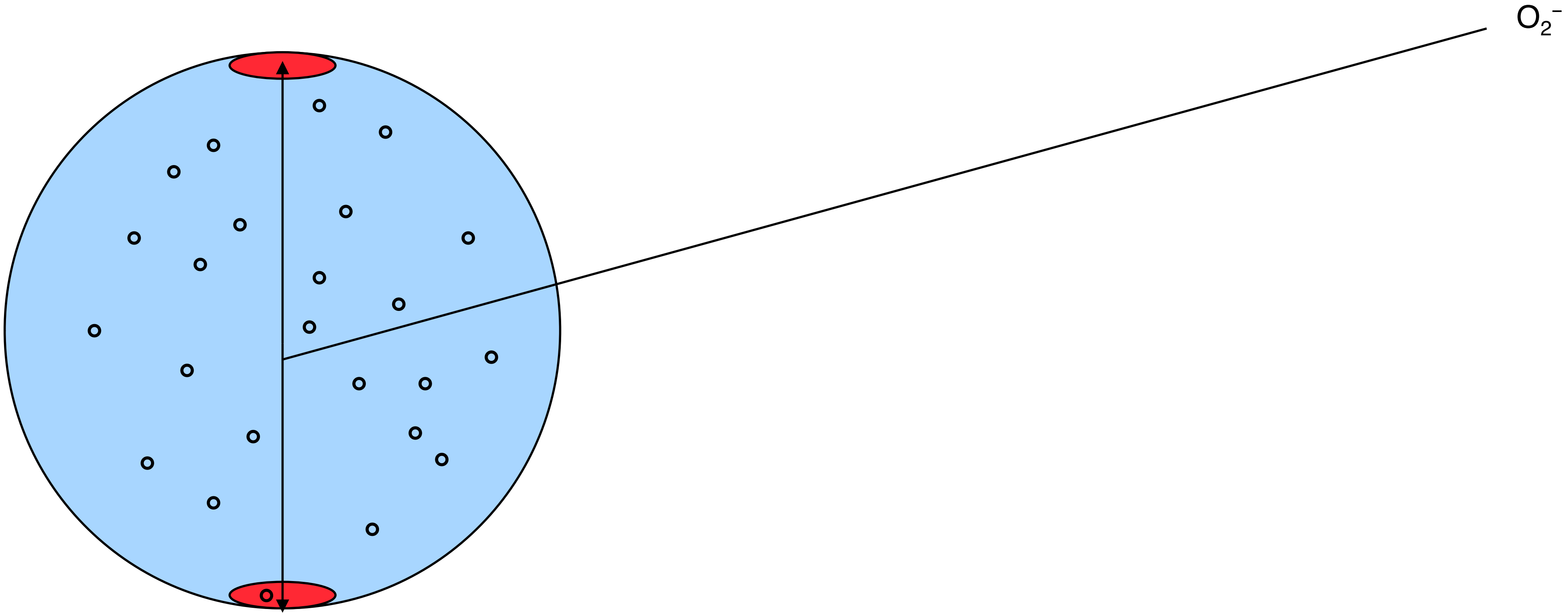
Electrostatic Model of SOD System

The earliest molecular recognition events are controlled by long-range forces.



Electrostatic Model of SOD System

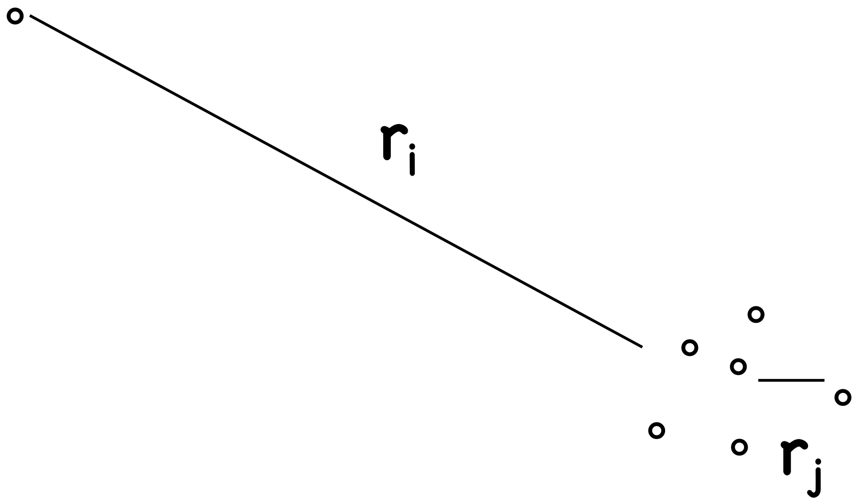
The earliest molecular recognition events are controlled by long-range forces. The electrostatic potential due to 76 charges of SOD can be described as a multipole expansion about the origin:



Multipole Expansion

The electrostatic potential energy at some point r_i due to a distribution of charges near some origin can be described as a multipole expansion

$$V_{elect} = q_i \sum_j \frac{q_j}{\left| \vec{r}_i - \vec{r}_j \right|}$$

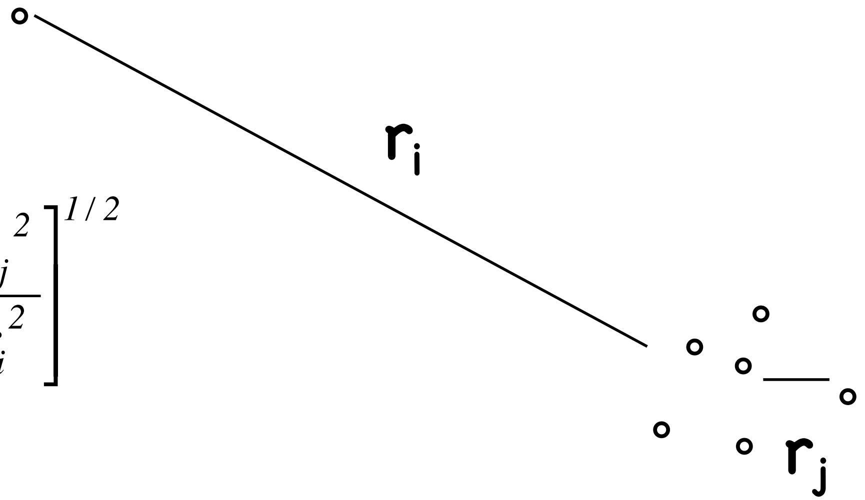


Multipole Expansion

The electrostatic potential energy at some point r_i due to a distribution of charges near some origin can be described as a multipole expansion

$$V_{elect} = q_i \sum_j \frac{q_j}{\left|\vec{r}_i - \vec{r}_j\right|}$$

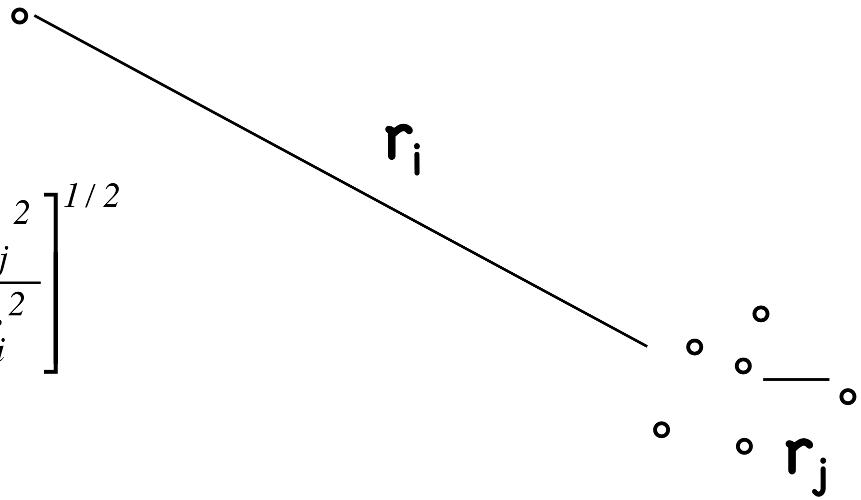
$$\left|\vec{r}_i - \vec{r}_j\right| = \left|r_i^2 - 2\vec{r}_i \cdot \vec{r}_j + r_j^2\right|^{1/2} = r_i\left[1 - \frac{2\vec{r}_i \cdot \vec{r}_j}{r_i} + \frac{r_j{}^2}{r_i{}^2}\right]^{1/2}$$



Multipole Expansion

The electrostatic potential energy at some point r_i due to a distribution of charges near some origin can be described as a multipole expansion

$$V_{elect} = q_i \sum_j \frac{q_j}{\left|\vec{r}_i - \vec{r}_j\right|}$$



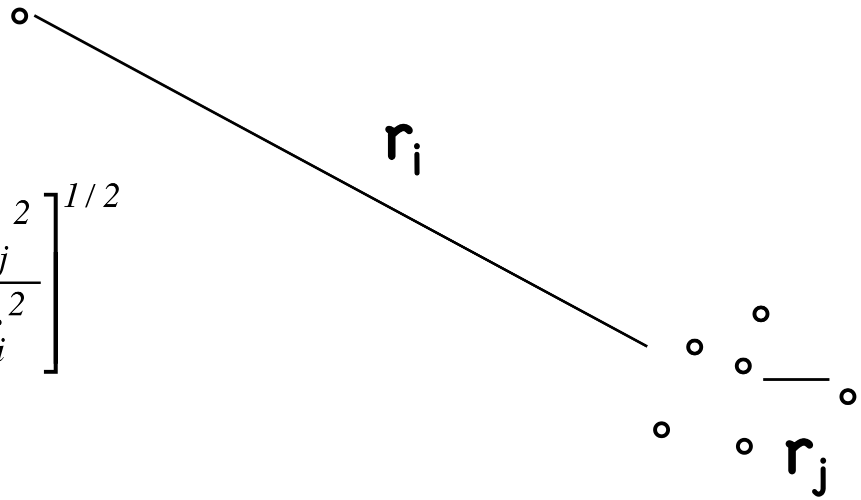
$$\left|\vec{r}_i - \vec{r}_j\right| = \left|r_i^2 - 2\vec{r}_i \cdot \vec{r}_j + r_j^2\right|^{1/2} = r_i\left[1 - \frac{2\vec{r}_i \cdot \vec{r}_j}{r_i^2} + \frac{r_j^{\ 2}}{r_i^{\ 2}}\right]^{1/2}$$

$$\frac{1}{\left|\vec{r}_i - \vec{r}_j\right|} = \frac{1}{r_i\left[1 - \frac{2\vec{r}_i \cdot \vec{r}_j}{r_i^2} + \frac{r_j^{\ 2}}{r_i^{\ 2}}\right]^{1/2}}$$

Multipole Expansion

The electrostatic potential energy at some point r_i due to a distribution of charges near some origin can be described as a multipole expansion

$$V_{elect} = q_i \sum_j \frac{q_j}{\left|\vec{r}_i - \vec{r}_j\right|}$$



$$\left|\vec{r}_i - \vec{r}_j\right| = \left|r_i^2 - 2\vec{r}_i \cdot \vec{r}_j + r_j^2\right|^{1/2} = r_i\left[1 - \frac{2\vec{r}_i \cdot \vec{r}_j}{r_i^2} + \frac{r_j^{\;2}}{r_i^{\;2}}\right]^{1/2}$$

$$\frac{1}{\left|\vec{r}_i - \vec{r}_j\right|} = \frac{1}{r_i\left[1 - \frac{2\vec{r}_i \cdot \vec{r}_j}{r_i^{\;2}} + \frac{r_j^{\;2}}{r_i^{\;2}}\right]^{1/2}}$$

Multipole Expansion

Since $r_j \ll r_i$, we can reexpress the 1/ri-rj as a binomial expansion (i.e x is small)

$$\left[1-\frac{2\vec{r}_i\cdot\vec{r}_j}{r_i^2}+\frac{r_j^{\,2}}{r_i^2}\right]^{-1/2}=\left[1-x\right]^{-1/2}=\left[1+\frac{1}{2}x+\frac{3}{8}x^2+\frac{15}{48}x^3+\ldots\right]$$

$$V_{elect}=q_i\sum_j\frac{q_j}{\left|\vec{r}_i-\vec{r}_j\right|}=\frac{q_i}{r_i}\sum_j q_j\left[1+\frac{\vec{r}_i\cdot\vec{r}_j}{r_i^2}-\frac{r_j^2}{2r_i^2}+\frac{3\left(\vec{r}_i\cdot\vec{r}_j\right)^2}{2r_i^4}-\frac{3r_j^2\left(\vec{r}_i\cdot\vec{r}_j\right)}{2r_i^4}+\frac{3r_j^4}{8r_i^4}\ldots\right]$$

$$V_{elect}=q_i\sum_j\frac{q_j}{\left|\vec{r}_i-\vec{r}_j\right|}=\frac{q_i}{r_i}\sum_j q_j\left[1+\frac{\vec{r}_i\cdot\vec{r}_j}{r_i^2}+\frac{3\left(\vec{r}_i\cdot\vec{r}_j\right)^2}{2r_i^4}-\frac{r_j^2}{2r_i^2}\right]$$

Since $r_j \ll r_i$, last 2 terms are ignored

Multipole Expansion

Rexpress previous equation:

$$V_{elect} = q_i \sum_j \frac{q_j}{\left|\vec{r}_i - \vec{r}_j\right|} = \frac{q_i}{r_i} \sum_j q_j \left[1 + \frac{\vec{r}_i \cdot \vec{r}_j}{r_i^2} + \frac{1}{2r_i^4} \left[3\left(\vec{r}_i \cdot \vec{r}_j\right)^2 - \left(\vec{r}_i \cdot \vec{r}_i\right)\left(\vec{r}_j \cdot \vec{r}_j\right) \right] \right]$$

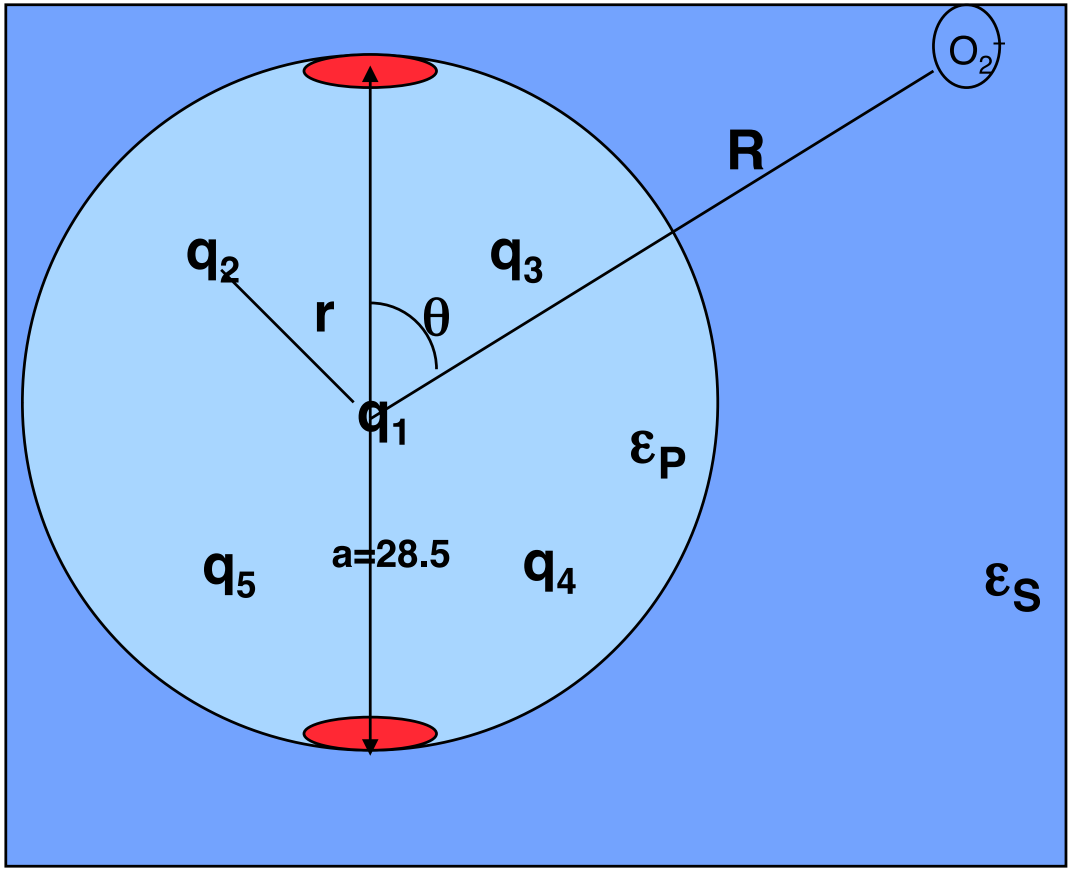
Rexpress this equation in terms of vector components:

$$V_{elect} = q_i \sum_j \frac{q_j}{\left|\vec{r}_i - \vec{r}_j\right|} = \frac{q_i}{r_i} \sum_j q_j \left[1 + \frac{\vec{r}_i \cdot \vec{r}_j}{r_i^2} + \frac{1}{2r_i^4} \sum_{\alpha=1}^{3} \sum_{\beta=1}^{3} \left[r_{i\alpha} r_{i\beta} \left(3 r_{j\alpha} r_{j\beta} - \delta_{\alpha\beta} r_j^2 \right) \right] \right]$$

Which gives us following multipole coefficients:

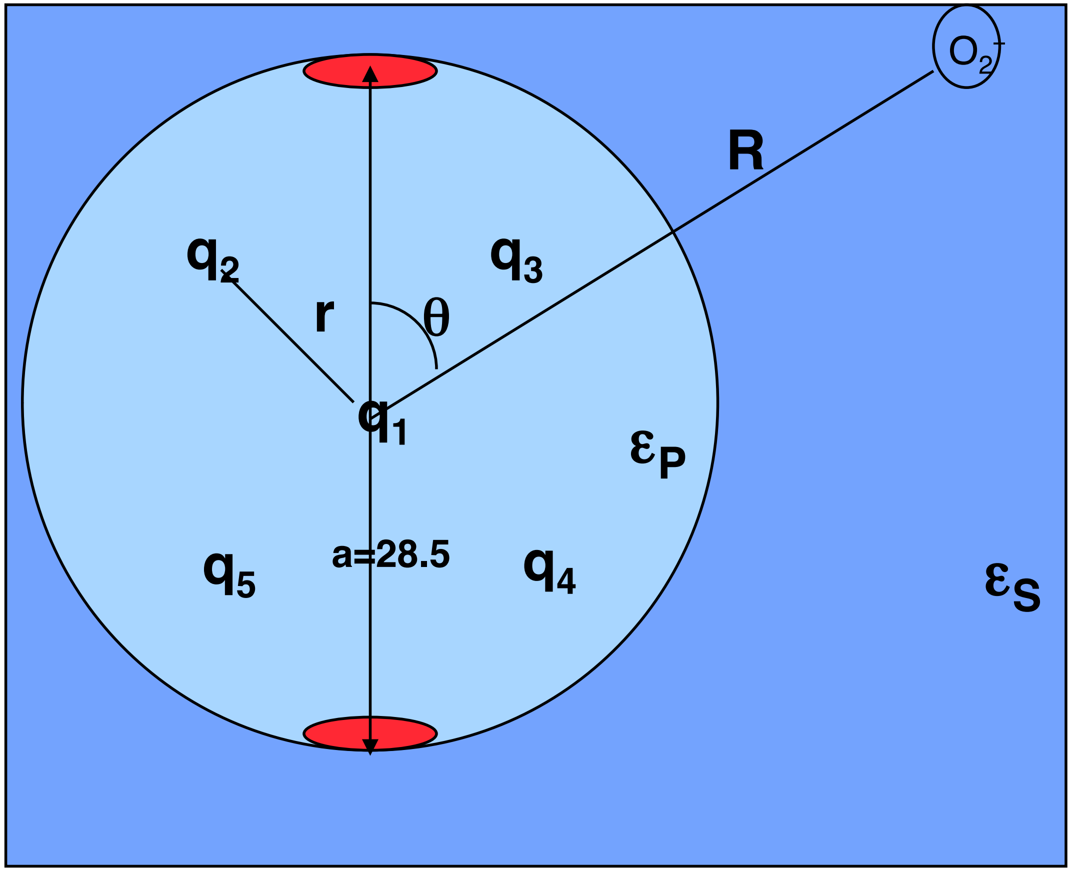
$$Q_M = \sum_j q_j \qquad Q_D = \sum_j q_j \vec{r}_i \cdot \vec{r}_j \qquad Q_{Q_{\alpha\beta}} = \sum_j \sum_\alpha \sum_\beta q_j \left[3 r_{j\alpha} r_{j\beta} - \delta_{\alpha\beta} r_j^2 \right]$$

Coarse-Graining Electrostatics



O_2^- is small sphere of q=-1

Coarse-Graining Electrostatics

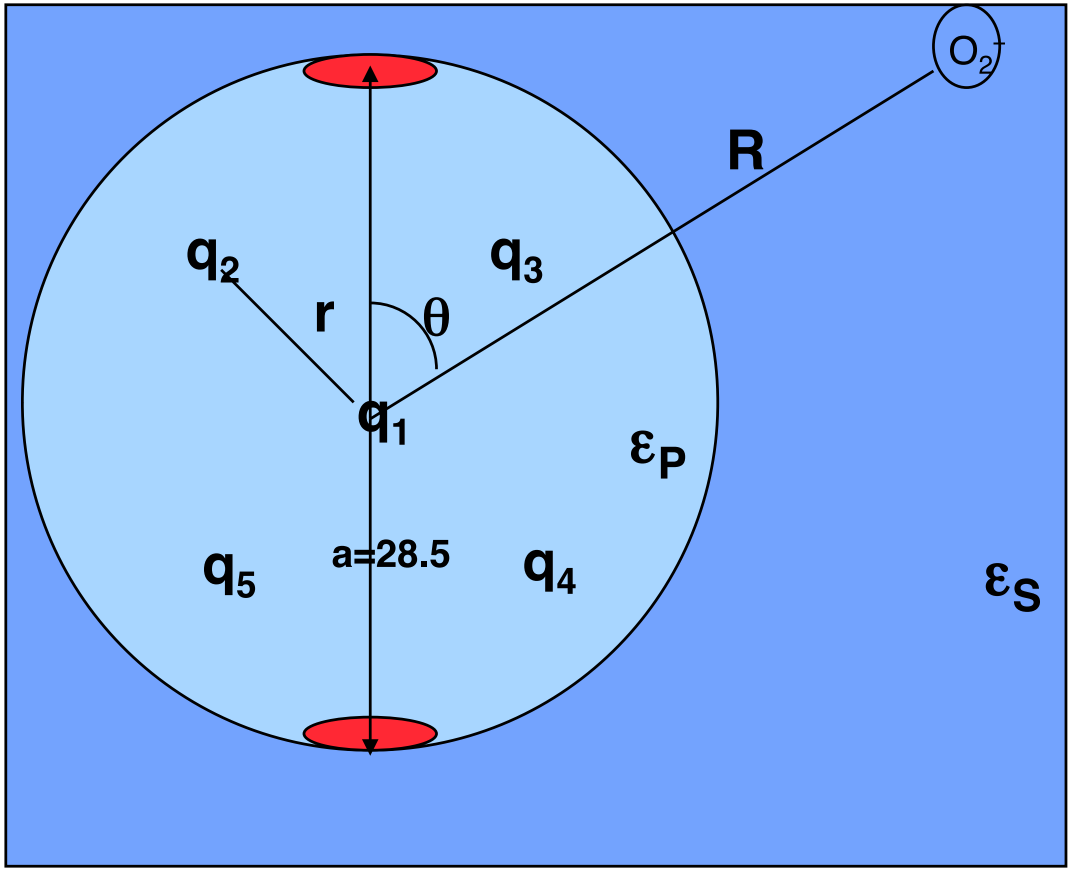


O_2^- is small sphere of q=-1

The 76 charges of SOD are reduced to a multipole expansion about the origin

Monopole=-4e
Dipole~0
Quadrupole: large

Coarse-Graining Electrostatics



O_2^- is small sphere of q=-1

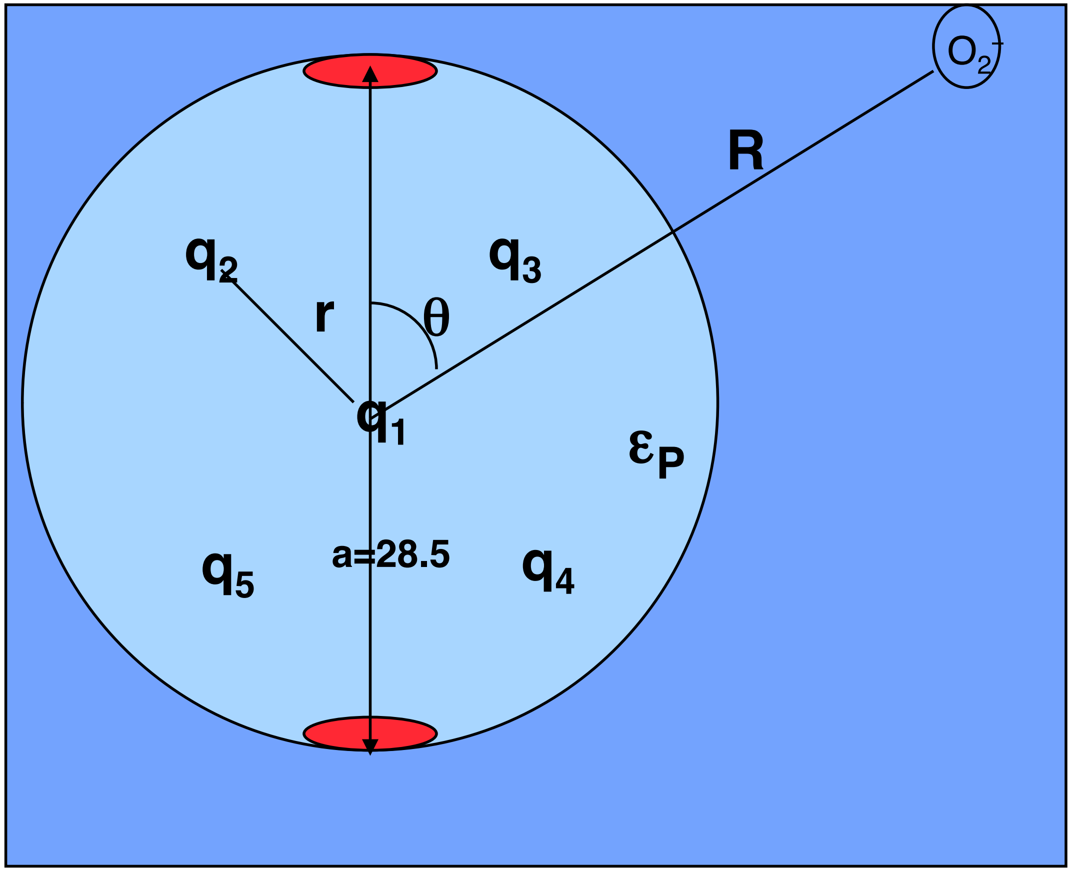
The 76 charges of SOD are reduced to a multipole expansion about the origin

Monopole=-4e
Dipole~0
Quadrupole: large

Now just place five charges in such a way that there multipole expansion reproduces same moments:

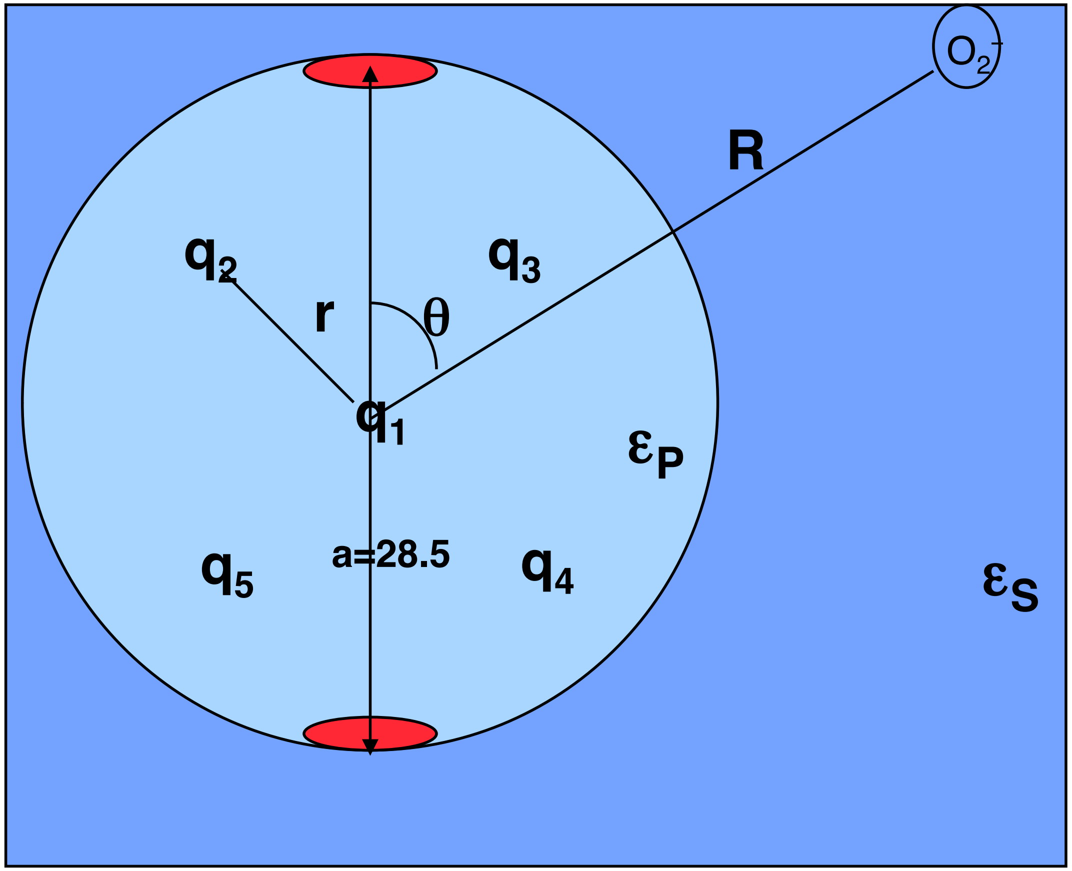
This is equivalent to 5 charges arranged as shown

Coarse-Grained Model of SOD System



Charges are enclosed in sphere of r=28.5A of low dielectric (ε_P)

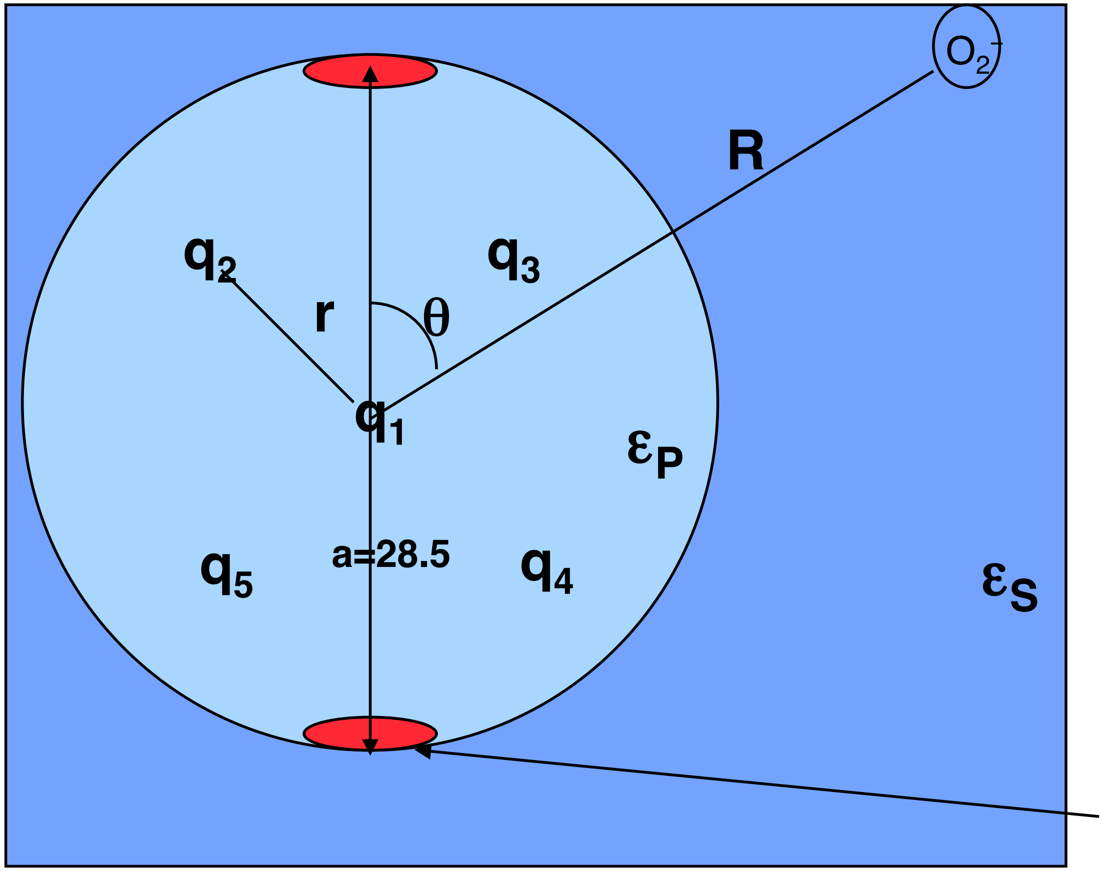
Coarse-Grained Model of SOD System



Charges are enclosed in sphere of r=28.5A of low dielectric (ε_P)

Outside protein is salty aqueous environment of high dielectric (ε_s)

Coarse-Grained Model of SOD System

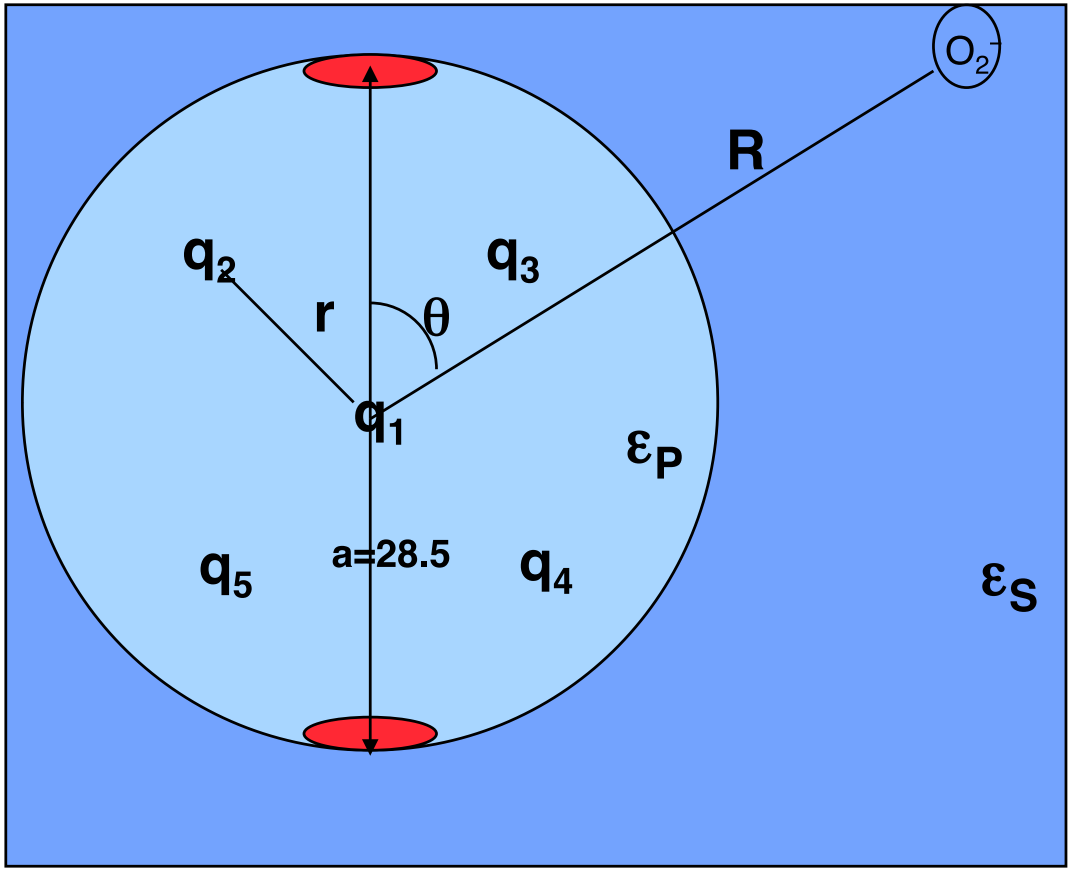


Charges are enclosed in sphere of r=28.5A of low dielectric (ε_P)

Outside protein is salty aqueous environment of high dielectric (ε_s)

Active sites of dimer are two caps which occupy 1/150th of protein surface area

Coarse-Grained Model of SOD System



Charges are enclosed in sphere of r=28.5A of low dielectric (ε_P)

Outside protein is salty aqueous environment of high dielectric (ε_s)

Active sites of dimer are two caps which occupy 1/150th of protein surface area

How to model the salty solution? More explicit ions? Since they diffuse around, we are only interested in their average spatial distributions. Use Debye Huckel theory

The Poisson-Boltzmann Equation

Poisson's Equation: $$-\nabla\left[\varepsilon(\mathbf{r})\nabla\Phi(\mathbf{r})\right]=\rho(\mathbf{r})=\rho_{fixed}(\mathbf{r})+\rho_{ions}(\mathbf{r})$$

Mean Field approximation: $$\rho_{ions}(\mathbf{r})=\sum_{j=1}^{K}c_j q_j\exp\left(-q_j\Phi(\mathbf{r})/kT\right)$$

Non-linear PB $$-\nabla\left[\varepsilon(\mathbf{r})\nabla\Phi(\mathbf{r})\right]+\underbrace{\sum_{j=1}^{K}c_j q_j\exp\left(-q_j\Phi(\mathbf{r})/kT\right)}_{\text{If 1:1 salt; } q\Phi/kT \ll 1}=\rho_{fixed}(\mathbf{r})$$

Linearized PB $$\boxed{-\nabla\left[\varepsilon(\mathbf{r})\nabla\Phi(\mathbf{r})\right]+\kappa^2\Phi(\mathbf{r})=\rho_{fixed}(\mathbf{r})}$$

Inverse Debye Length $$\kappa=\sqrt{\frac{8\pi e^2 C}{\varepsilon_s kT}}$$

Debye-Huckel Theory

What do solutions of the mean field equations look like?

Debye-Huckel Theory

What do solutions of the mean field equations look like?

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\Phi}{dr}\right) = -\frac{4\pi}{\varepsilon}\rho(r)$$

Poissons equation in spherical coordinates

Debye-Huckel Theory

What do solutions of the mean field equations look like?

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\Phi}{dr}\right) = -\frac{4\pi}{\varepsilon}\rho(r)$$

Poissons equation in spherical coordinates

$$\int_a^\infty 4\pi r^2\rho(r)dr = 0$$

We will enforce electroneutrality

Debye-Huckel Theory

What do solutions of the mean field equations look like?

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\Phi}{dr}\right) = -\frac{4\pi}{\varepsilon}\rho(r)$$

Poissons equation in spherical coordinates

$$\int_a^\infty 4\pi r^2 \rho(r)dr = 0$$

We will enforce electroneutrality

$$n_i' = n_i \exp(-q_i\Phi / kT)$$

Ions are distributed according to Boltzmann distribution with interaction energy $-q_i\Phi(r)$

Debye-Huckel Theory

What do solutions of the mean field equations look like?

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\Phi}{dr}\right) = -\frac{4\pi}{\varepsilon}\rho(r)$$

Poissons equation in spherical coordinates

$$\int_a^\infty 4\pi r^2 \rho(r) dr = 0$$

We will enforce electroneutrality

$$n_i' = n_i \exp\left(-q_i\Phi / kT\right)$$

Ions are distributed according to Boltzmann distribution with interaction energy $-q_i\Phi(r)$

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\Phi}{dr}\right) = -\frac{4\pi}{\varepsilon}\sum_i q_i n_i \exp\left(-q_i\Phi / kT\right)$$

Density=$\Sigma q_i n_i'$

$$\exp\left(-q_i\Phi / kT\right) = 1 - q_i\Phi / kT$$

Debye-Huckel Theory

What do solutions of the mean field equations look like?

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\Phi}{dr}\right) = -\frac{4\pi}{\varepsilon}\rho(r)$$

Poissons equation in spherical coordinates

$$\int_a^\infty 4\pi r^2\rho(r)dr = 0$$

We will enforce electroneutrality

$$n_i' = n_i\,exp\left(-q_i\Phi / kT\right)$$

Ions are distributed according to Boltzmann distribution with interaction energy $-q_i\Phi(r)$

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\Phi}{dr}\right) = -\frac{4\pi}{\varepsilon}\sum_i q_i n_i\,exp\left(-q_i\Phi / kT\right)$$

Density=$\Sigma q_i n_i'$

$$exp\left(-q_i\Phi / kT\right) = 1 - q_i\Phi / kT$$

For low [monovalent salts], Φ<<kT and do expansion of exponential

Debye-Huckel Theory

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\Phi}{dr}\right) = -\cancelto{0}{\frac{4\pi}{\varepsilon}\sum_i q_i n_i} + \frac{4\pi}{\varepsilon}\sum_i q_i^2 n_i \Phi_j / kT$$

Debye-Huckel Theory

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\Phi}{dr}\right) = -\cancelto{0}{\frac{4\pi}{\varepsilon}\sum_i q_i n_i} + \frac{4\pi}{\varepsilon}\sum_i q_i^2 n_i \Phi_j / kT$$

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\Phi}{dr}\right) = \frac{4\pi}{\varepsilon}\sum_i q_i^2 n_i \Phi_j / kT$$

Debye-Huckel Theory

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\Phi}{dr}\right) = -\frac{4\pi}{\varepsilon}\cancelto{0}{\sum_i q_i n_i} + \frac{4\pi}{\varepsilon}\sum_i q_i^2 n_i \Phi_j / kT$$

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\Phi}{dr}\right) = \frac{4\pi}{\varepsilon}\sum_i q_i^2 n_i \Phi_j / kT$$

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\Phi}{dr}\right) = \kappa^2 \Phi_j$$

κ is the DH screening length. It is what we will use to control salt concentration in the model.

Debye-Huckel Theory

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\Phi}{dr}\right) = -\cancelto{0}{\frac{4\pi}{\varepsilon}\sum_i q_i n_i} + \frac{4\pi}{\varepsilon}\sum_i q_i^2 n_i \Phi_j / kT$$

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\Phi}{dr}\right) = \frac{4\pi}{\varepsilon}\sum_i q_i^2 n_i \Phi_j / kT$$

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\Phi}{dr}\right) = \kappa^2 \Phi_j$$

κ is the DH screening length. It is what we will use to control salt concentration in the model.

$$\Phi_j = \frac{A}{R} exp(\kappa R)$$

With a solution that will look something like this…

Electrostatic Potential Energy

The general form of the potential for charges inside low dielectric sphere and for outside high dielectric solvent

$$\Psi_{in} = \sum_{n=1}^{\#charges} \sum_{m=1} \left(\frac{q_n r^m}{R^{m+1}} + B_m R^m \right) P_m(\cos\theta) \qquad 0 < R < a$$

$$\Psi_{out} = \sum_{m=1} \frac{A_m \exp(-\kappa R) K_m(\kappa R) P_m(\cos\theta)}{R^{m+1}} \qquad R > a$$

Electrostatic Potential Energy

The general form of the potential for charges inside low dielectric sphere and for outside high dielectric solvent

$$\Psi_{in} = \sum_{n=1}^{\#\,charges} \sum_{m=1} \left(\frac{q_n r^m}{R^{m+1}} + B_m R^m \right) P_m(\cos\theta) \qquad 0 < R < a$$

$$\Psi_{out} = \sum_{m=1} \frac{A_m \exp(-\kappa R) K_m(\kappa R) P_m(\cos\theta)}{R^{m+1}} \qquad R > a$$

$P_m(\cos\theta)$ Legendre polynomials

$K_m(\kappa R)$ Modified spherical bessel functions

Electrostatic Potential Energy

The unknown coefficients Am and Bm are solved to give the final form of the potential by enforcing boundary conditions. These are the requirements that the potential must be continuous across the dielectric boundary:

Electrostatic Potential Energy

The unknown coefficients Am and Bm are solved to give the final form of the potential by enforcing boundary conditions. These are the requirements that the potential must be continuous across the dielectric boundary:

$$-\frac{1}{a}\frac{d\Psi_{in}}{d\theta}\bigg|_{R=a} = -\frac{1}{a}\frac{d\Psi_{out}}{d\theta}\bigg|_{R=a}$$

Tangential to boundary

$$-\varepsilon_{SOD}\frac{d\Psi_{in}}{dR}\bigg|_{R=a} = -\varepsilon_{out}\frac{d\Psi_{out}}{dR}\bigg|_{R=a}$$

Normal to boundary

Electrostatic Potential Energy

The unknown coefficients Am and Bm are solved to give the final form of the potential by enforcing boundary conditions. These are the requirements that the potential must be continuous across the dielectric boundary:

$$-\frac{1}{a}\frac{d\Psi_{in}}{d\theta}\bigg|_{R=a} = -\frac{1}{a}\frac{d\Psi_{out}}{d\theta}\bigg|_{R=a}$$

Tangential to boundary

$$-\varepsilon_{SOD}\frac{d\Psi_{in}}{dR}\bigg|_{R=a} = -\varepsilon_{out}\frac{d\Psi_{out}}{dR}\bigg|_{R=a}$$

Normal to boundary

$$\Psi_{out} = \sum_{n=1}^{\#charges}\sum_{m=1}\frac{qr^m(2m+1)K_m(\kappa R)P_m(\cos\theta)\exp(-\kappa(R-a))}{R^{m+1}\left[\varepsilon_{out}(m+1)K_{m+1}(\kappa a)+mK_m(\kappa a)(\varepsilon_{SOD}-\varepsilon_{out})\right]}$$

Electrostatic Potential Energy

The unknown coefficients Am and Bm are solved to give the final form of the potential by enforcing boundary conditions. These are the requirements that the potential must be continuous across the dielectric boundary:

$$-\frac{1}{a}\frac{d\Psi_{in}}{d\theta}\bigg|_{R=a} = -\frac{1}{a}\frac{d\Psi_{out}}{d\theta}\bigg|_{R=a}$$

Tangential to boundary

$$-\varepsilon_{SOD}\frac{d\Psi_{in}}{dR}\bigg|_{R=a} = -\varepsilon_{out}\frac{d\Psi_{out}}{dR}\bigg|_{R=a}$$

Normal to boundary

$$\Psi_{out} = \sum_{n=1}^{\#charges}\sum_{m=1}\frac{qr^m(2m+1)K_m(\kappa R)P_m(\cos\theta)\exp(-\kappa(R-a))}{R^{m+1}\left[\varepsilon_{out}(m+1)K_{m+1}(\kappa a)+mK_m(\kappa a)(\varepsilon_{SOD}-\varepsilon_{out})\right]}$$

and series expansions converges to m~10.

Brownian Motion: Brownian Dynamics

The potential:

$$\Psi_{out} = \sum_{n=1}^{\# charges} \sum_{m=1} \frac{q r^m (2m+1) K_m(\kappa R) P_m(\cos\theta) \exp(-\kappa(R-a))}{R^{m+1}\left[\varepsilon_{out}(m+1)K_{m+1}(\kappa a) + m K_m(\kappa a)(\varepsilon_{SOD} - \varepsilon_{out})\right]}$$

and derivative with respect to R can be defined to get force.

Brownian Motion: Brownian Dynamics

The potential:

$$\Psi_{out} = \sum_{n=1}^{\#charges} \sum_{m=1} \frac{qr^m(2m+1)K_m(\kappa R)P_m(\cos\theta)\exp(-\kappa(R-a))}{R^{m+1}\left[\varepsilon_{out}(m+1)K_{m+1}(\kappa a)+mK_m(\kappa a)(\varepsilon_{SOD}-\varepsilon_{out})\right]}$$

and derivative with respect to R can be defined to get force.

Solve the numerical BD equation of motion of O_2^- to SOD active sites

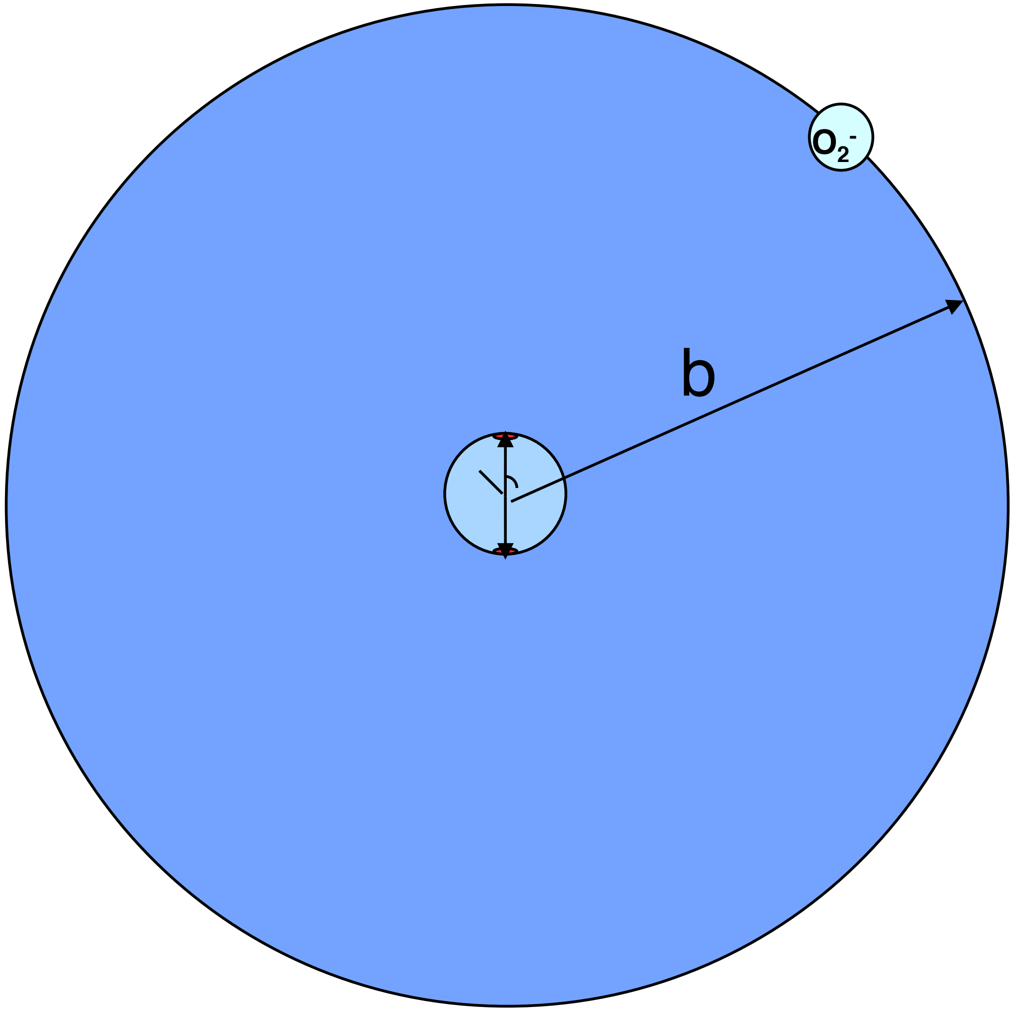
$$R(t+\Delta t) = R(t) + \frac{Dm}{kT} f(t)\Delta t + Q(t)$$

$$D = \frac{\left(a_{SOD} + a_{O_2^-}\right)kT}{4\pi\eta a_{SOD} a_{O_2^-}}$$

$$\langle Q(t)\rangle = 0 \qquad \left\langle Q(t)^2\right\rangle = 6D\Delta t$$

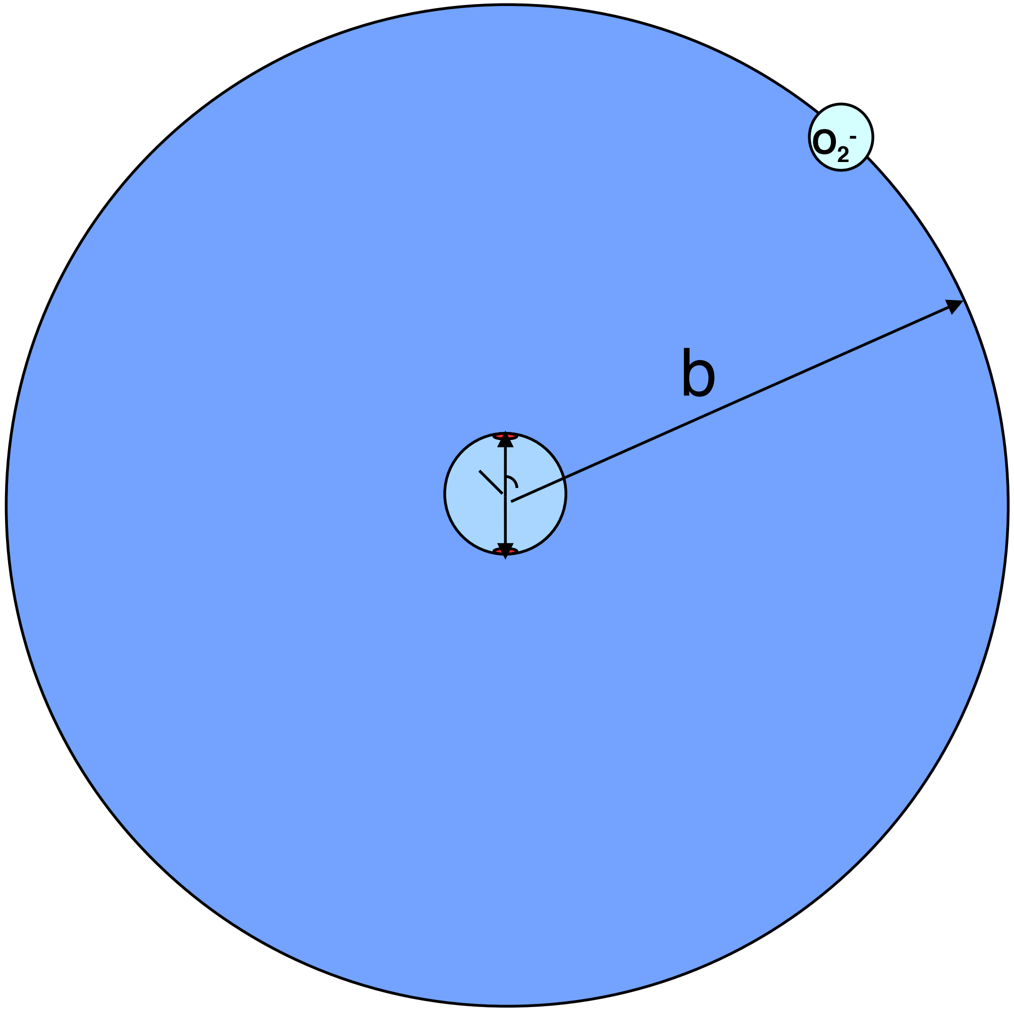
Simulating association rates of SOD/O_2^-

O_2^- starts at a distance b from the origin where the potential and therefore the flux at b is isotropic.



Simulating association rates of SOD/O_2^-

O_2^- starts at a distance b from the origin where the potential and therefore the flux at b is isotropic.

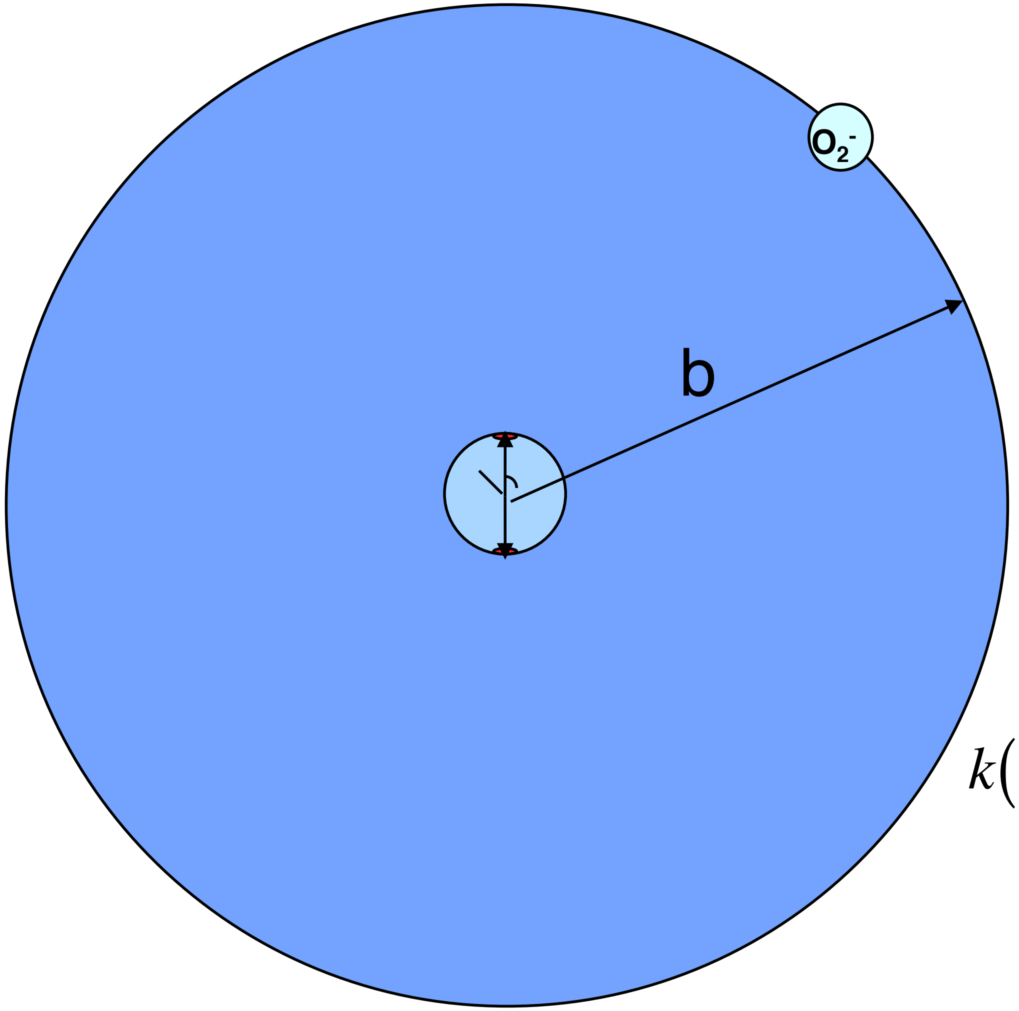


The rate can be expressed as

$$k=k(b)P$$

Simulating association rates of SOD/O_2^-

O_2^- starts at a distance b from the origin where the potential and therefore the flux at b is isotropic.



The rate can be expressed as

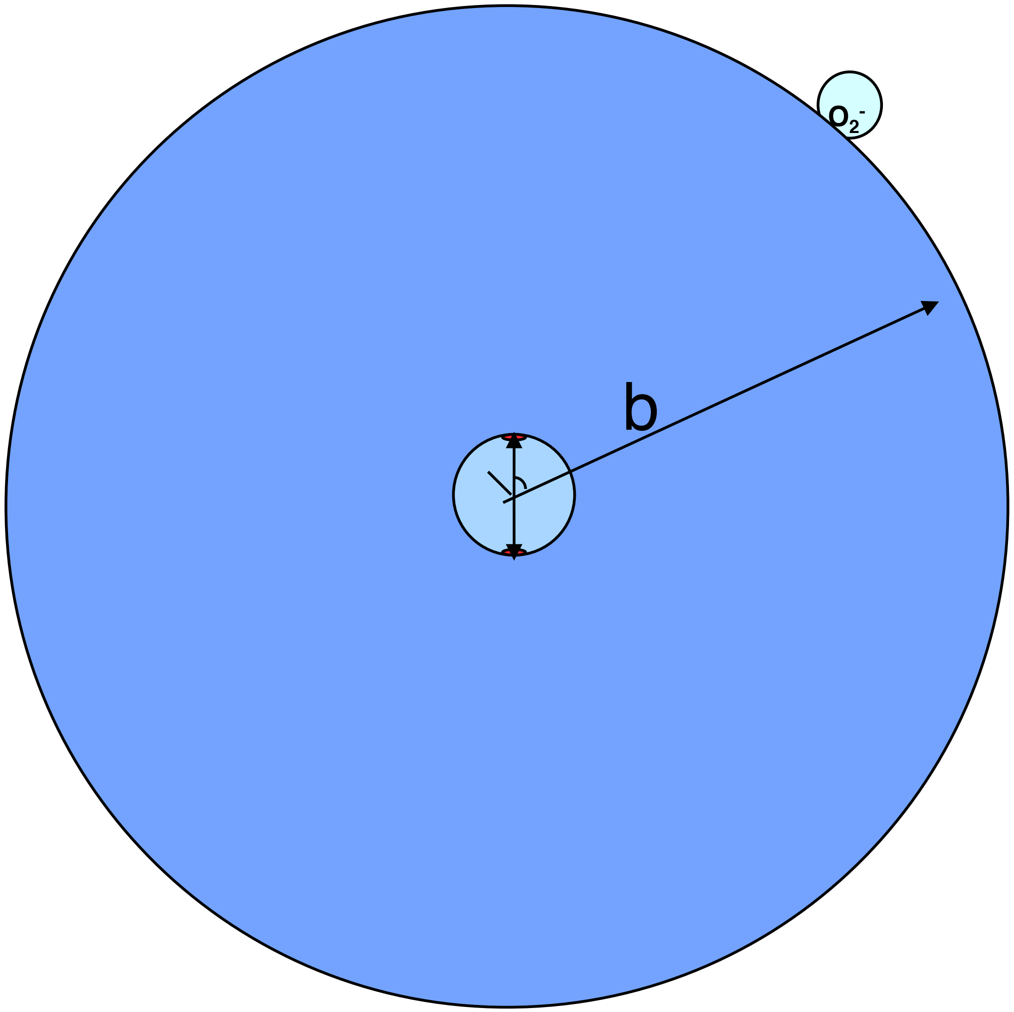
k=k(b)P

k(b) is rate for reaching b after diffusing from infinity and can be solved analytically

$$k(b) = \left[\frac{1}{4\pi D} \int_b^\infty \frac{exp\left(\beta q_{O_2^-} q_{SOD}^{monopole} / R\right)}{R^2} dR \right]^{-1}$$

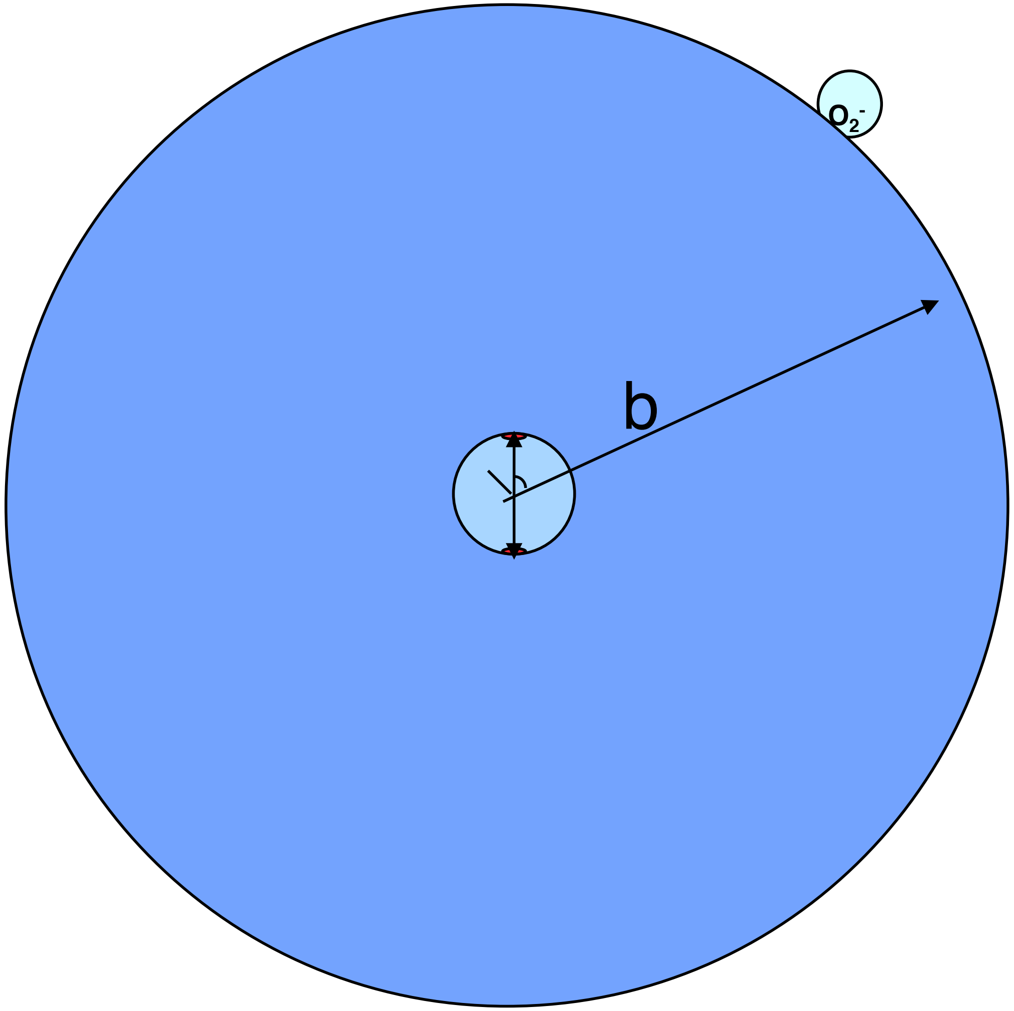
Simulating association rates of SOD/O_2^-

P = B_∞A is the probability that pair reacts and does not escape to infinity, A is activation energy of catalysis



Simulating association rates of SOD/O_2^-

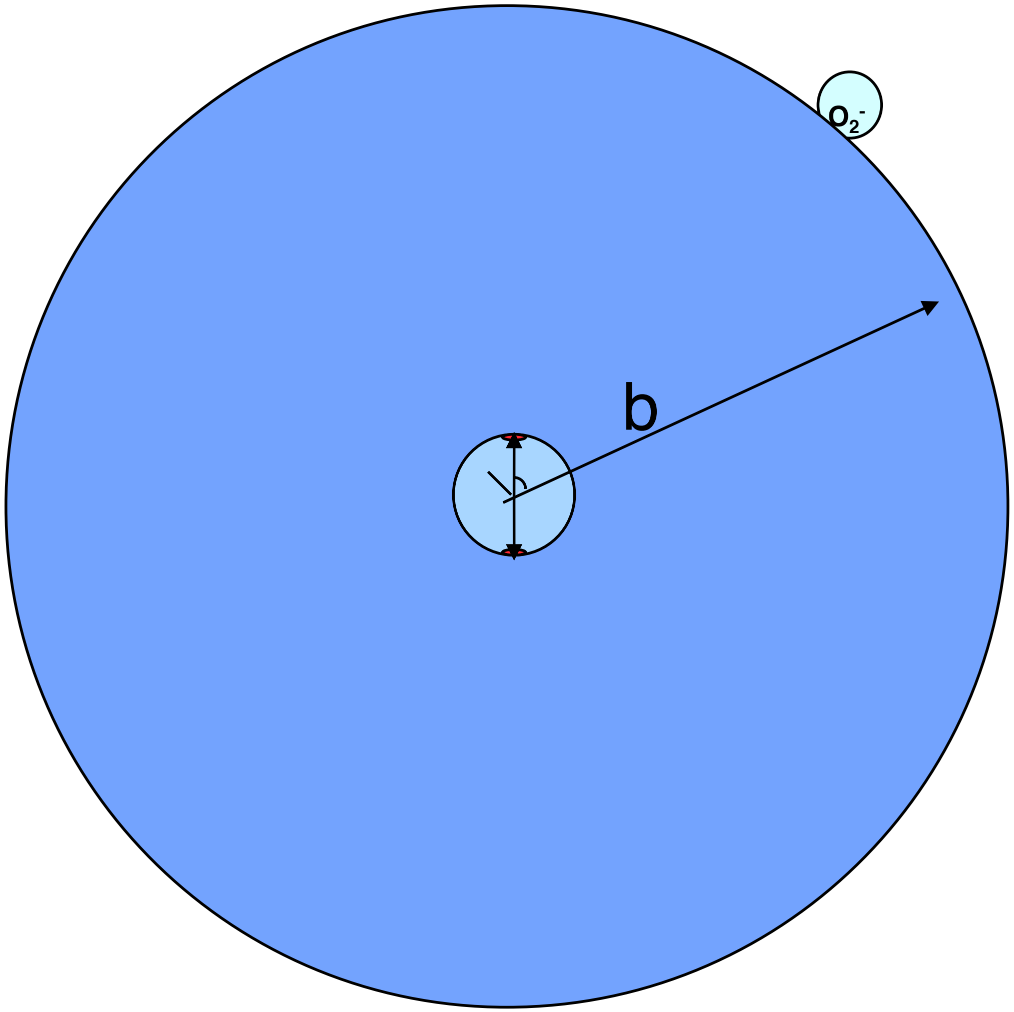
$P = B_\infty A$ is the probability that pair reacts and does not escape to infinity, A is activation energy of catalysis



Since experiment tells us this is diffusion-controlled we set A=1

Simulating association rates of SOD/O_2^-

$P = B_\infty A$ is the probability that pair reacts and does not escape to infinity, A is activation energy of catalysis

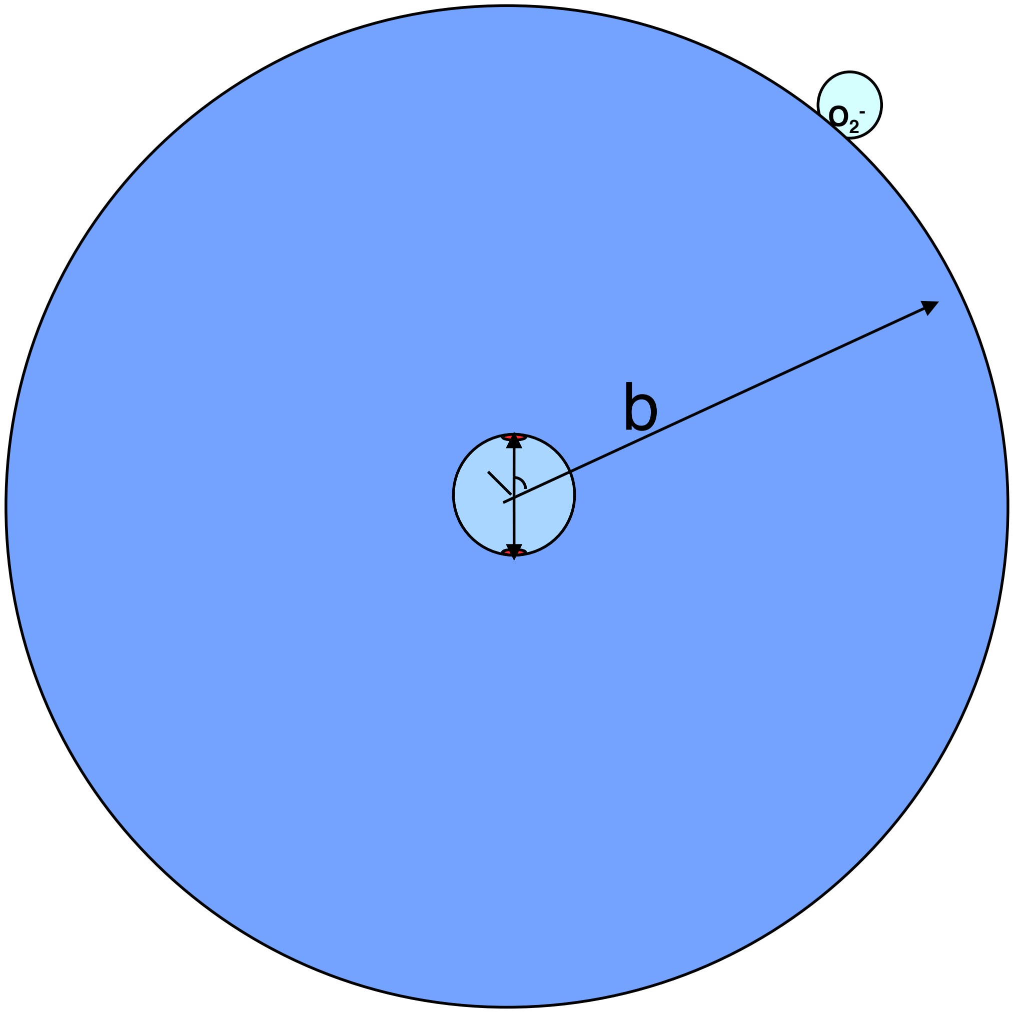


Since experiment tells us this is diffusion-controlled we set A=1

B_∞ is the probability that particle reaches b and makes it to the reactive surface when moving in an *infinite* domain.

Simulating association rates of SOD/O_2^-

$P = B_\infty A$ is the probability that pair reacts and does not escape to infinity, A is activation energy of catalysis



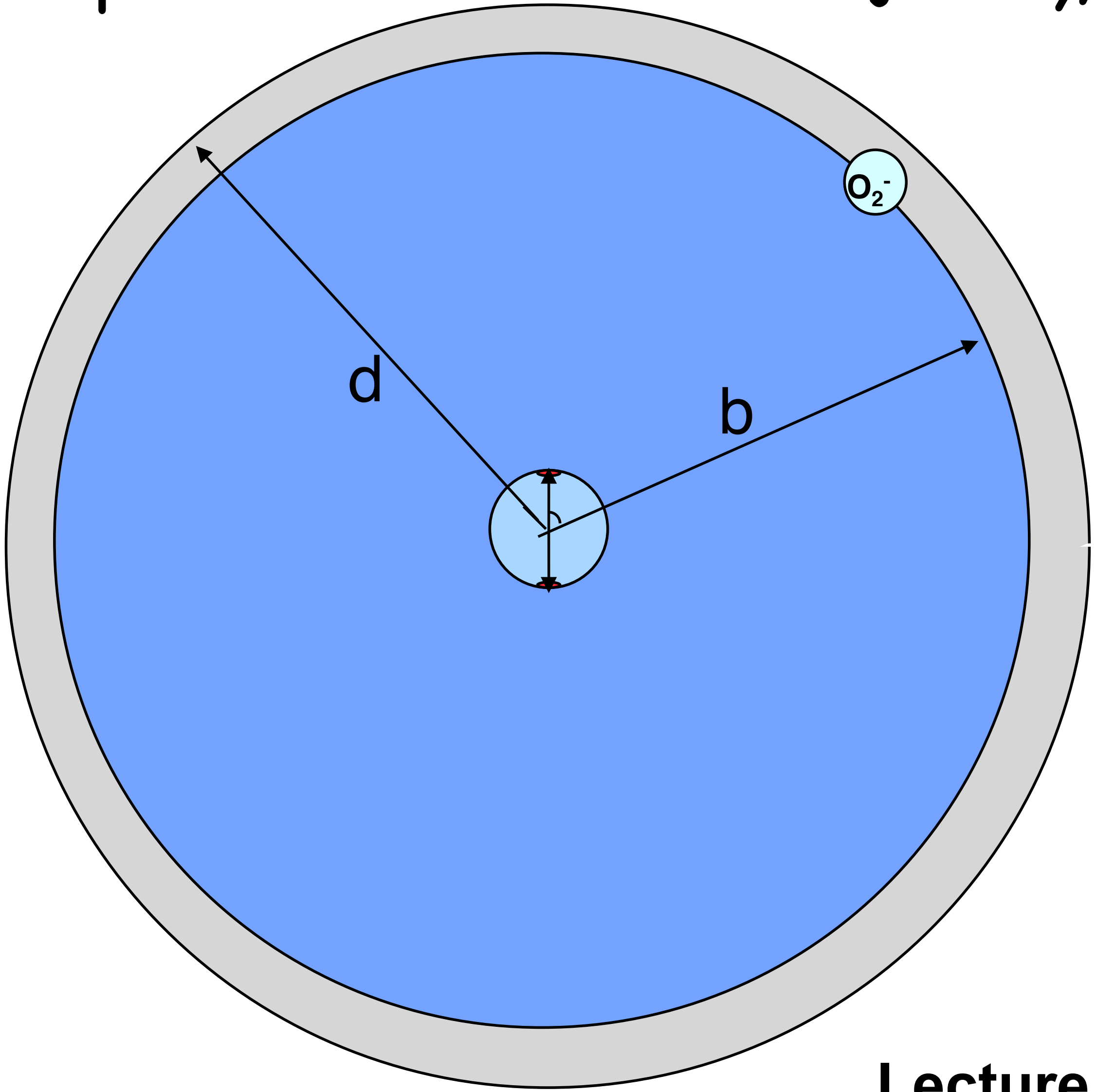
Since experiment tells us this is diffusion-controlled we set A=1

B_∞ is the probability that particle reaches b and makes it to the reactive surface when moving in an *infinite* domain.

But given all of these simplifications, it still would be impossible to simulate in an infinite domain.

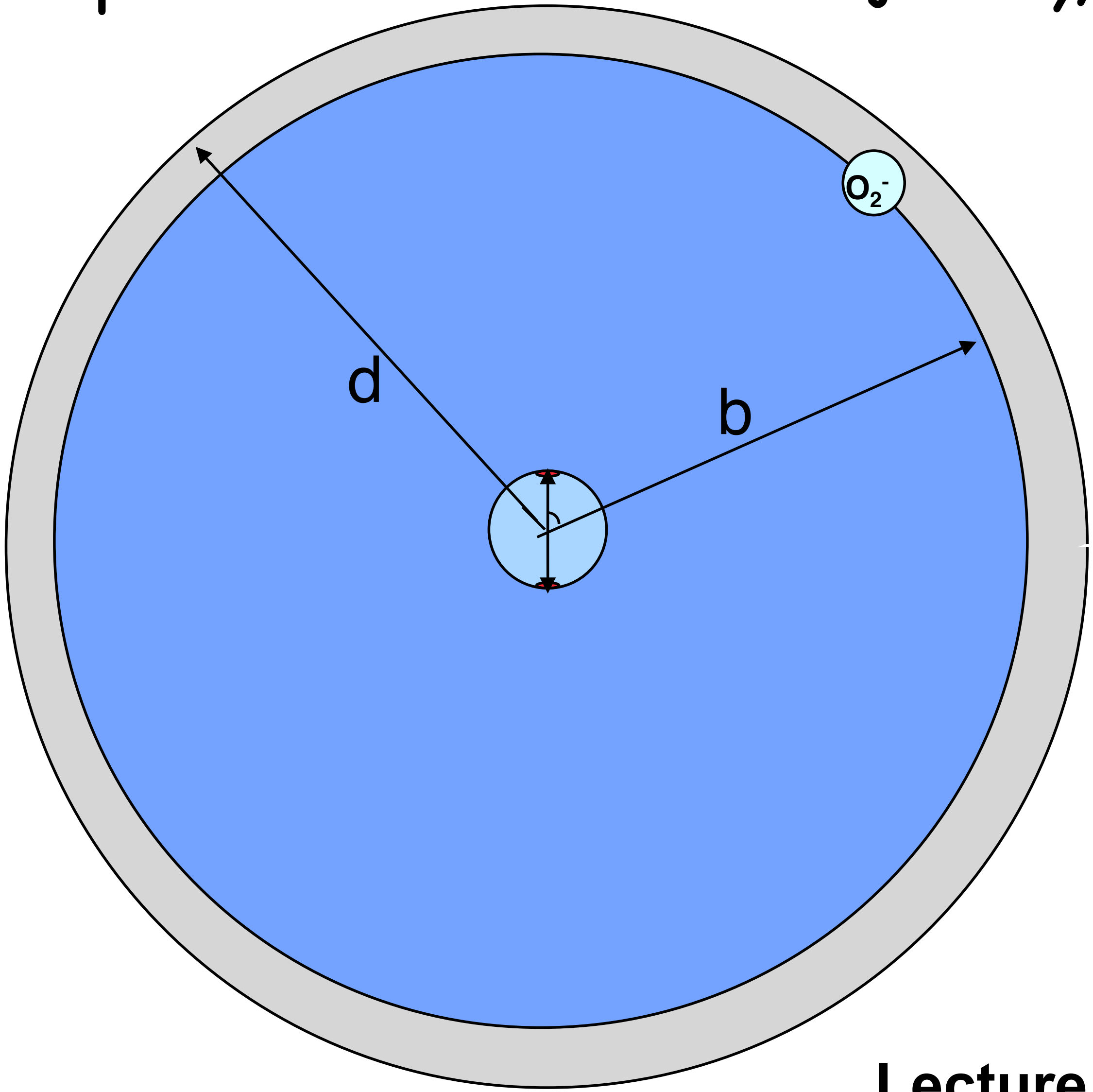
Simulating association rates of SOD/O_2^-

We define another sphere at d, and if O_2^- reaches this outer sphere we truncate the trajectory, and count the trajectory as non-reactive



Simulating association rates of SOD/O_2^-

We define another sphere at d, and if O_2^- reaches this outer sphere we truncate the trajectory, and count the trajectory as non-reactive



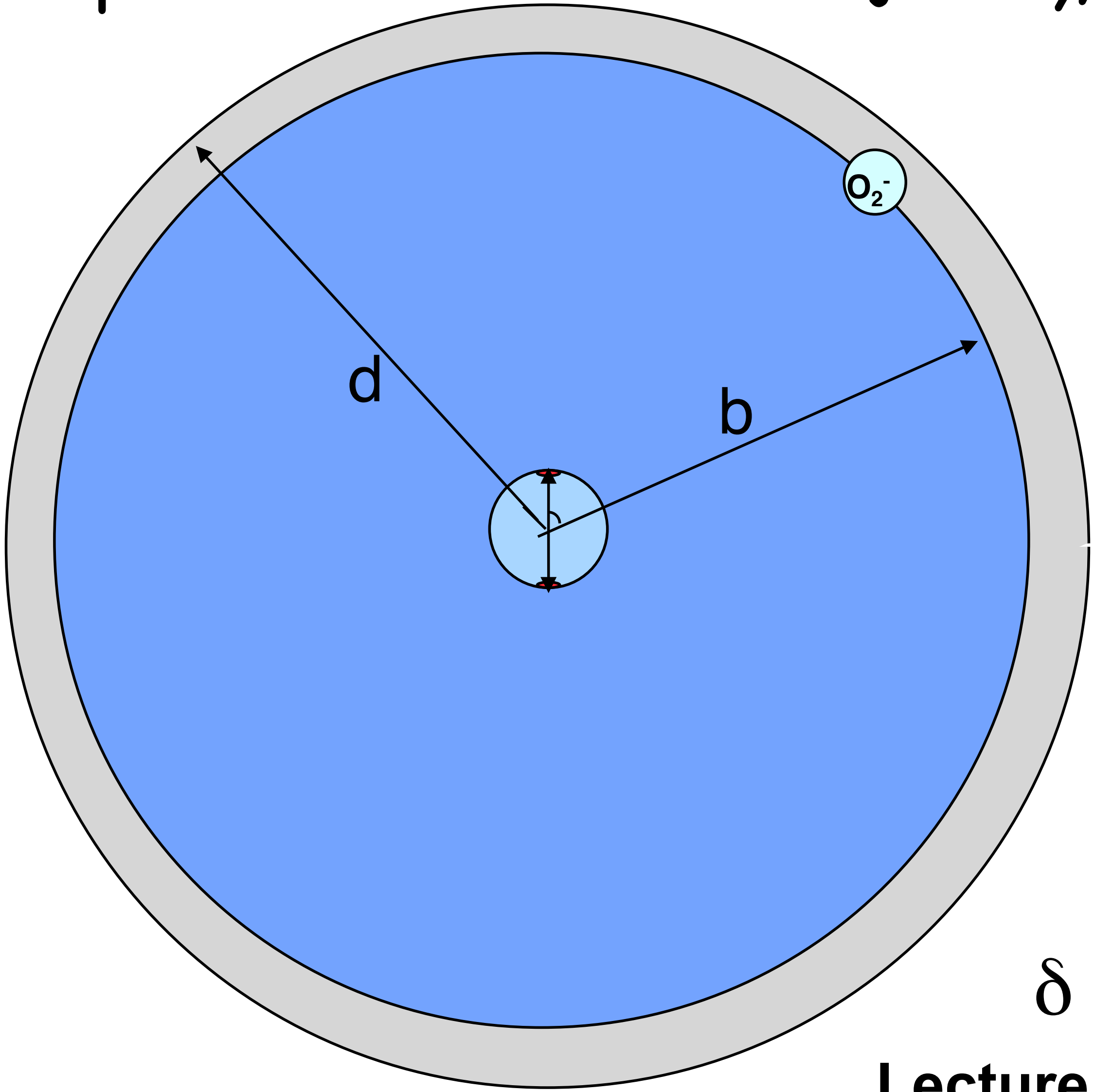
The probability B_∞ now becomes

$B_\infty \propto (1-\delta)$,

Where $(1-\delta)$ is the probability that particle reaches separation R=d.

Simulating association rates of SOD/O_2^-

We define another sphere at d, and if O_2^- reaches this outer sphere we truncate the trajectory, and count the trajectory as non-reactive



The probability B_∞ now becomes

$B_\infty \propto (1-\delta)$,

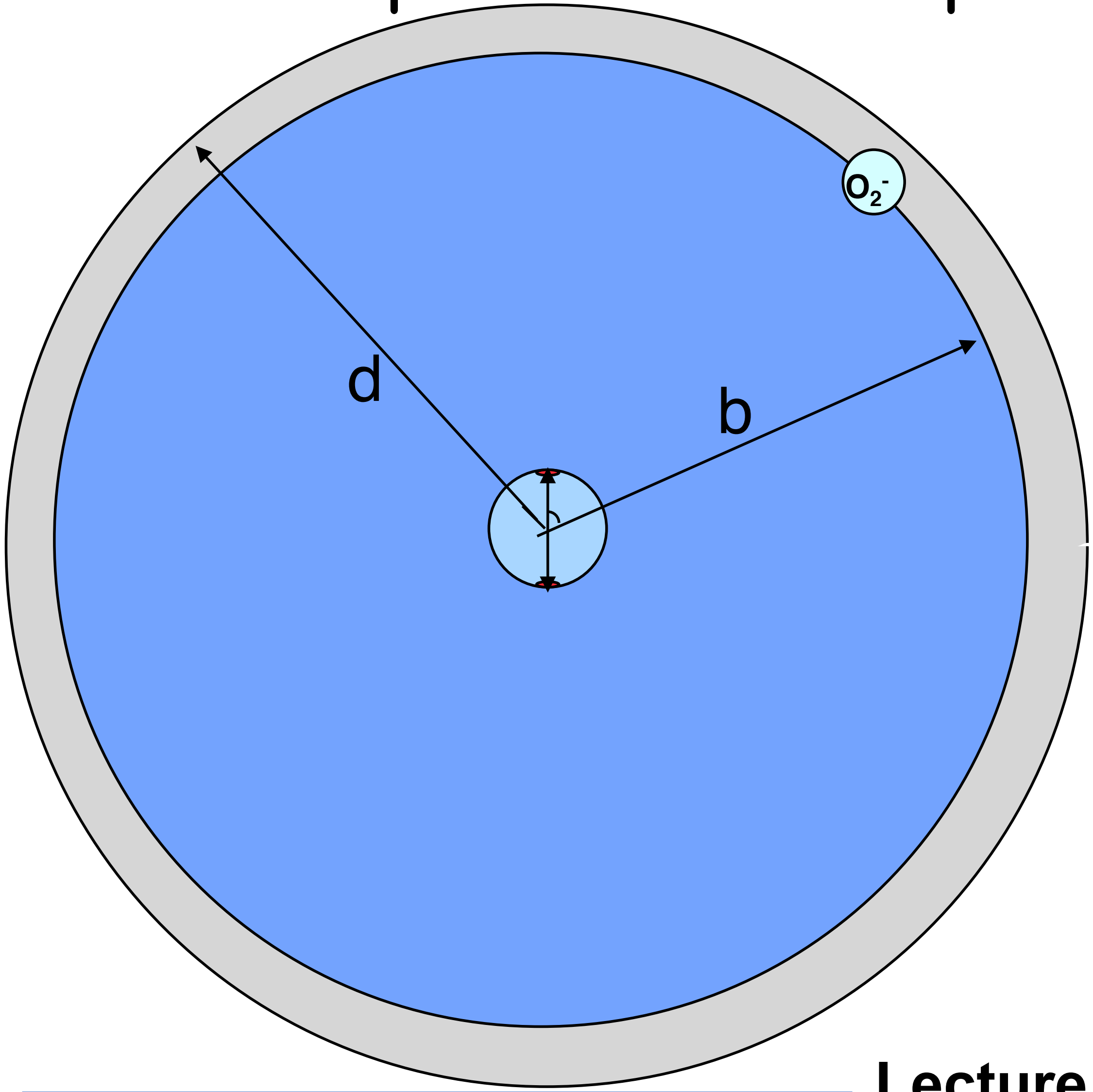
Where (1- δ) is the probability that particle reaches separation R=d.

δ is given by our simulation:

$$\delta = \frac{\text{\# trajectories reaching active site}}{\text{total \# trajectories}}$$

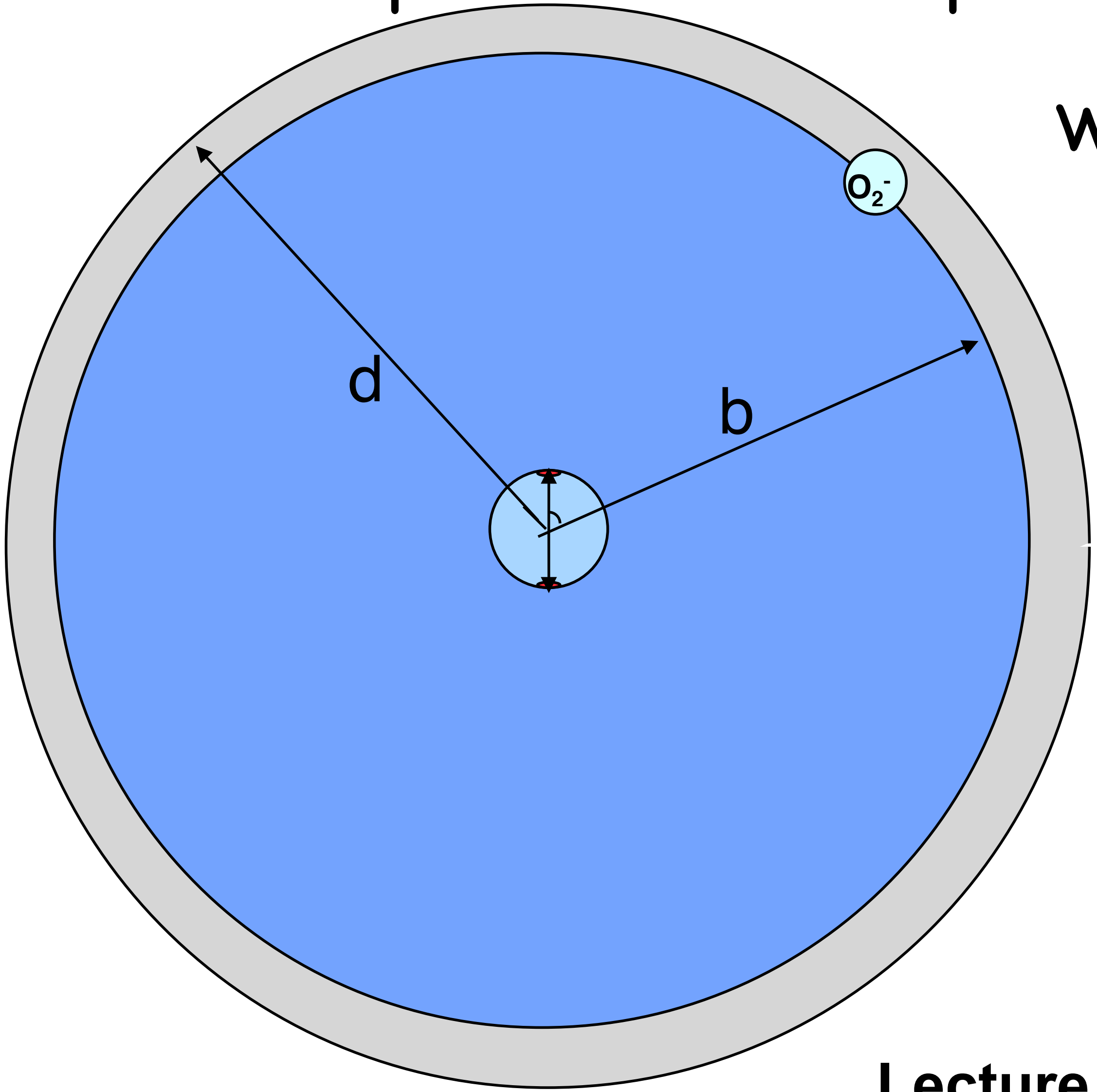
Simulating association rates of SOD/O_2^-

The probability B∞ is also proportional to (1-γ), the probability that if particle reaches separation R=d it escapes to infinity.



Simulating association rates of SOD/O_2^-

The probability B∞ is also proportional to (1-γ), the probability that if particle reaches separation R=d it escapes to infinity.

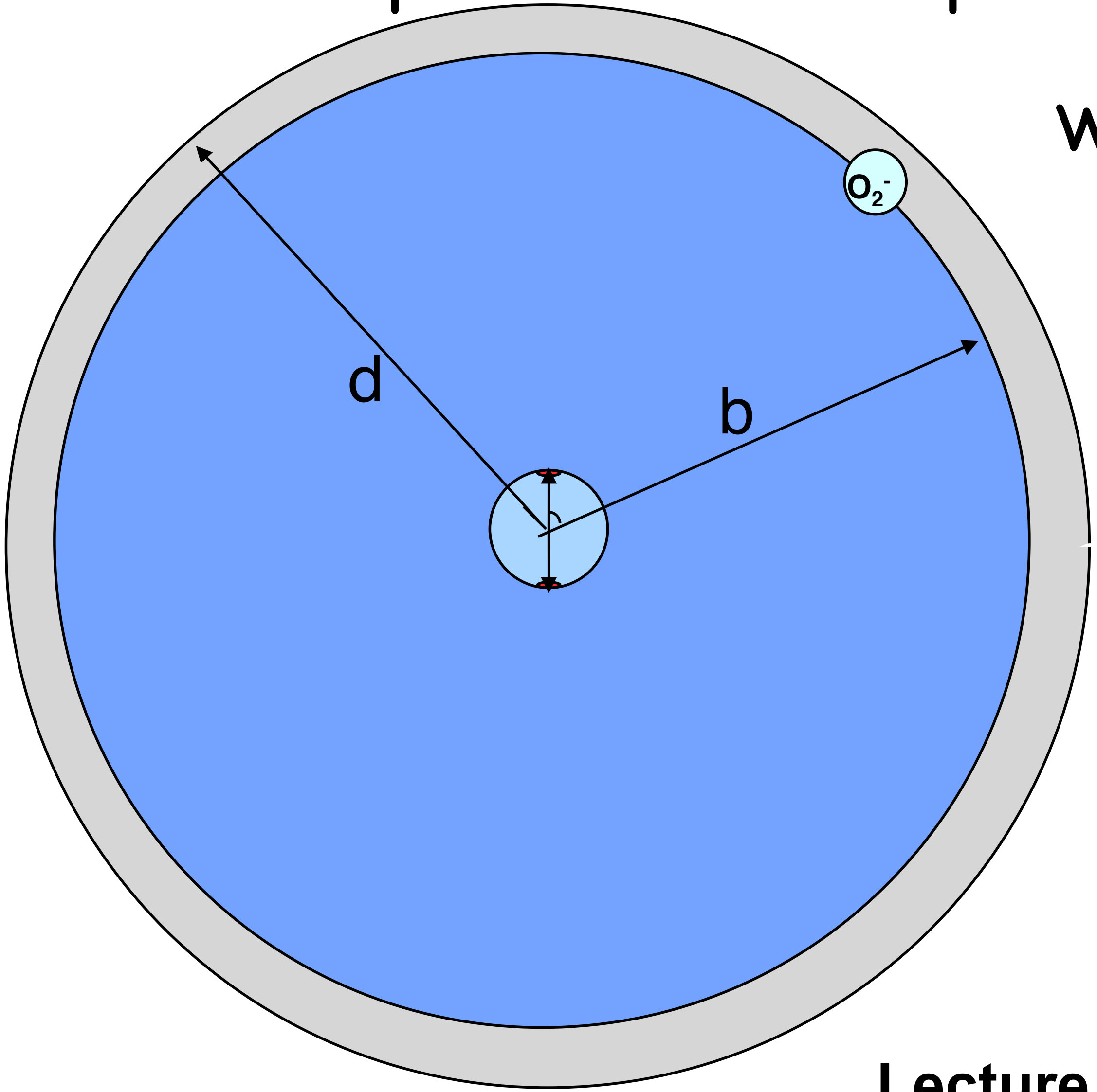


We can solve for γ analytically

$$\gamma = \frac{k(b)}{k(d)}$$

Simulating association rates of SOD/O_2^-

The probability B∞ is also proportional to (1-γ), the probability that if particle reaches separation R=d it escapes to infinity.



We can solve for γ analytically

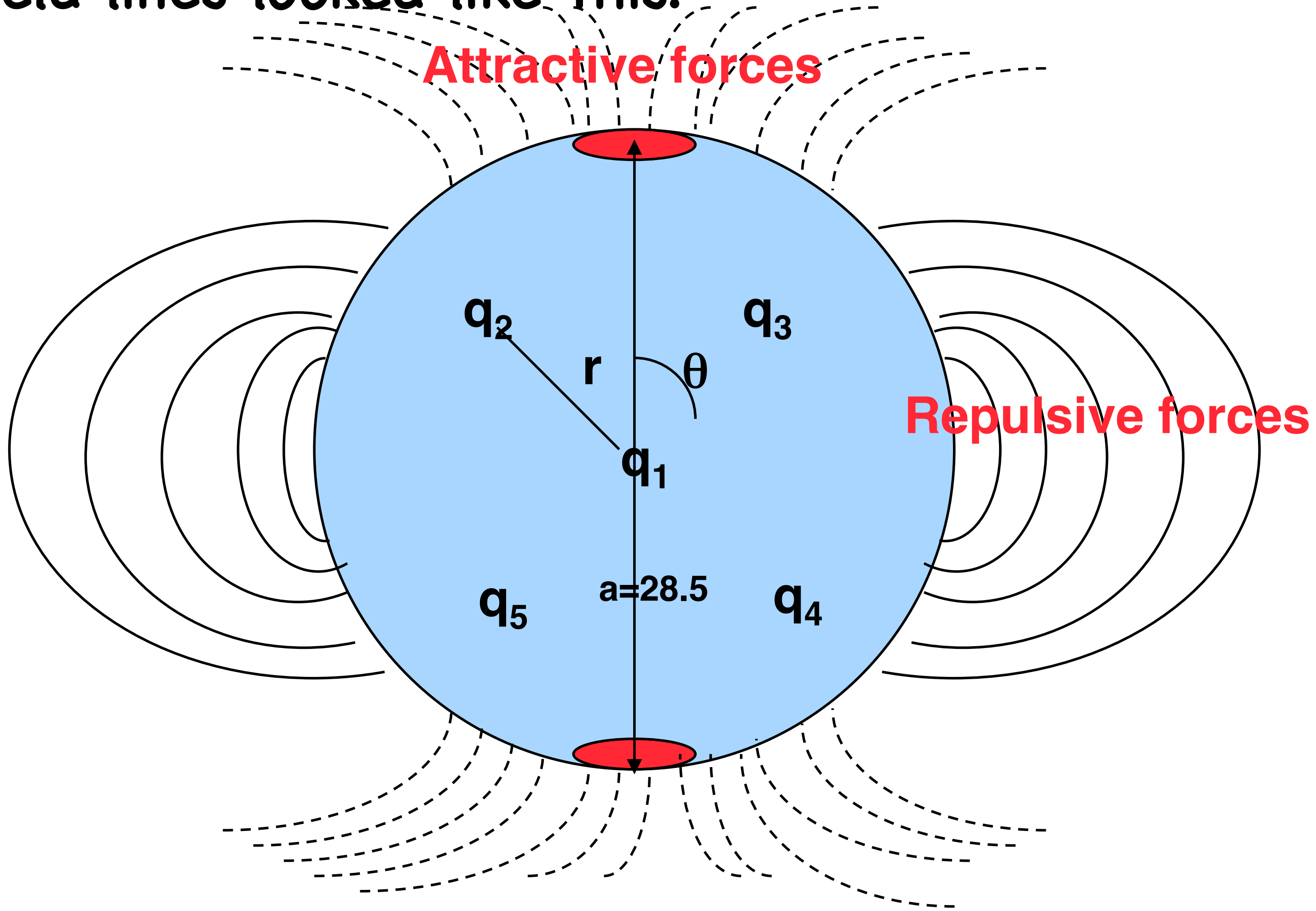
$$\gamma = \frac{k(b)}{k(q)}$$

However you can have multiple recrossings across the boundary d, so that final expression for B_∞.

$$B_\infty = \frac{\delta}{\left[1-(1-\delta)\gamma\right]}$$

Summary of Results for SOD/O_2^-

The electrostatic field lines looked like this:

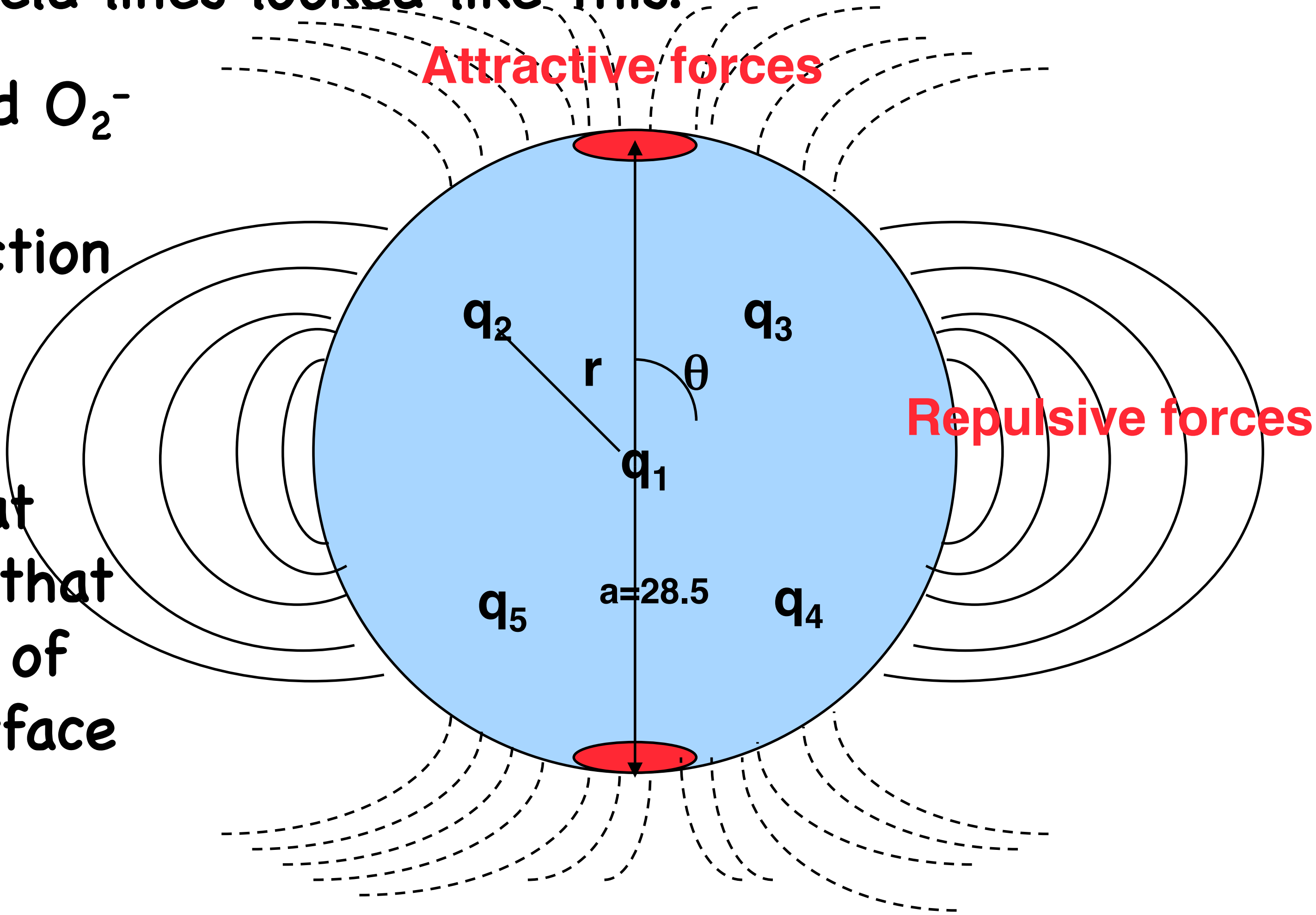


Summary of Results for SOD/O_2^-

The electrostatic field lines looked like this:

(1) Protein (-4e) and O_2^- (-1e) suggest a repulsive interaction that would slow down the rate.

(2) Active site sits at bottom of cleft that occupies 1/150th of total protein surface area.

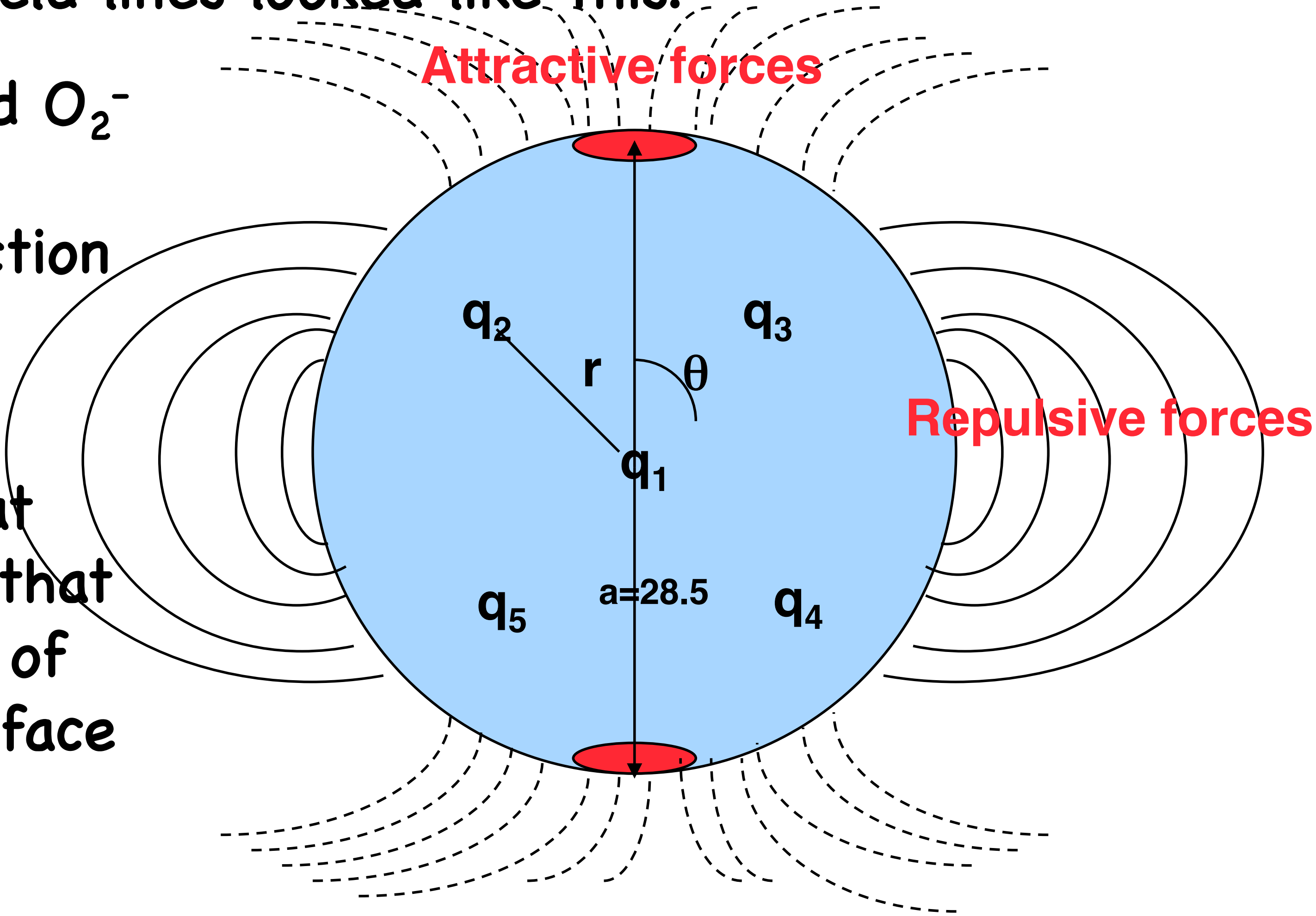


Summary of Results for SOD/O_2^-

The electrostatic field lines looked like this:

(1) Protein (−4e) and O_2^- (−1e) suggest a repulsive interaction that would slow down the rate.

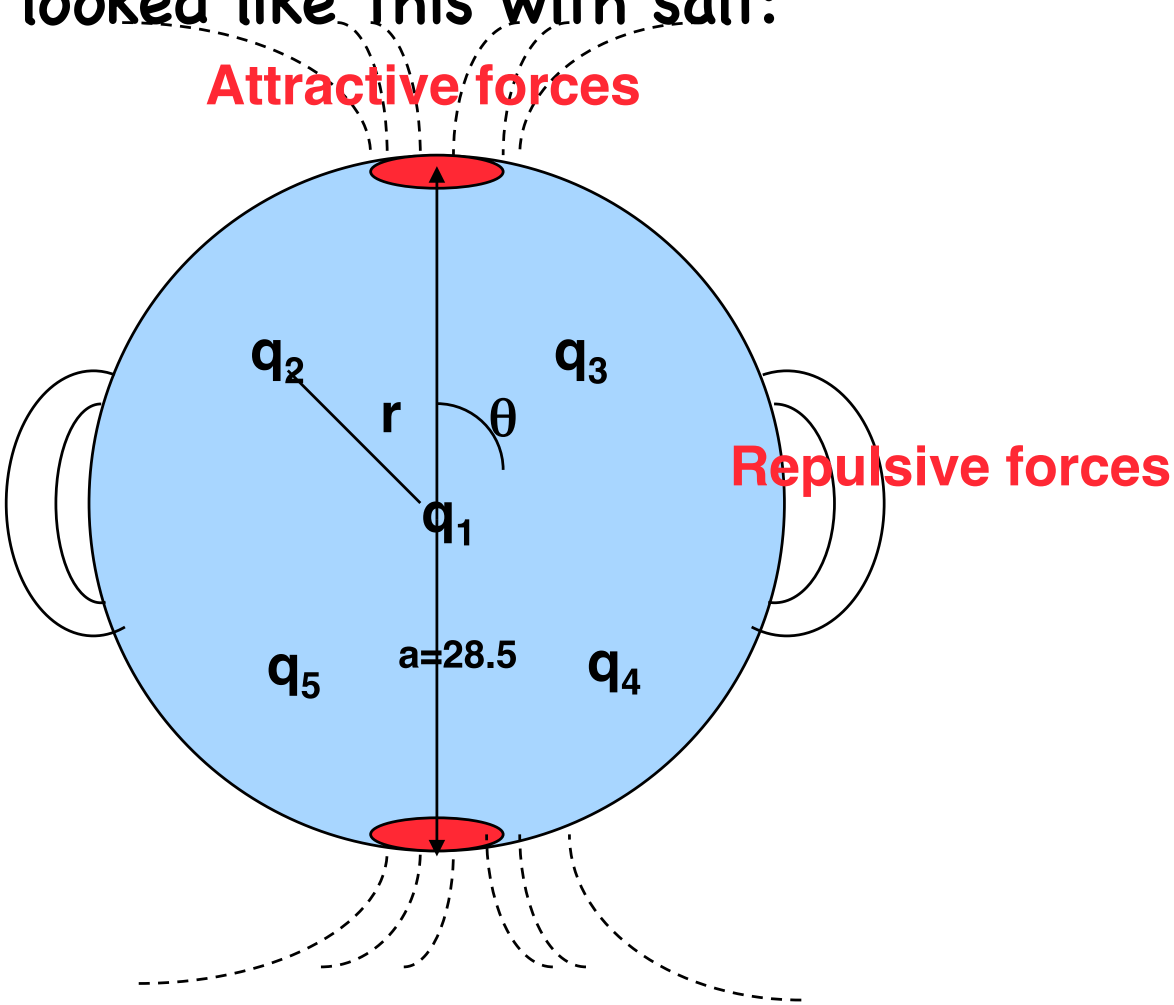
(2) Active site sits at bottom of cleft that occupies 1/150th of total protein surface area.



Effectively the potential steers superoxide toward the protein active sites.

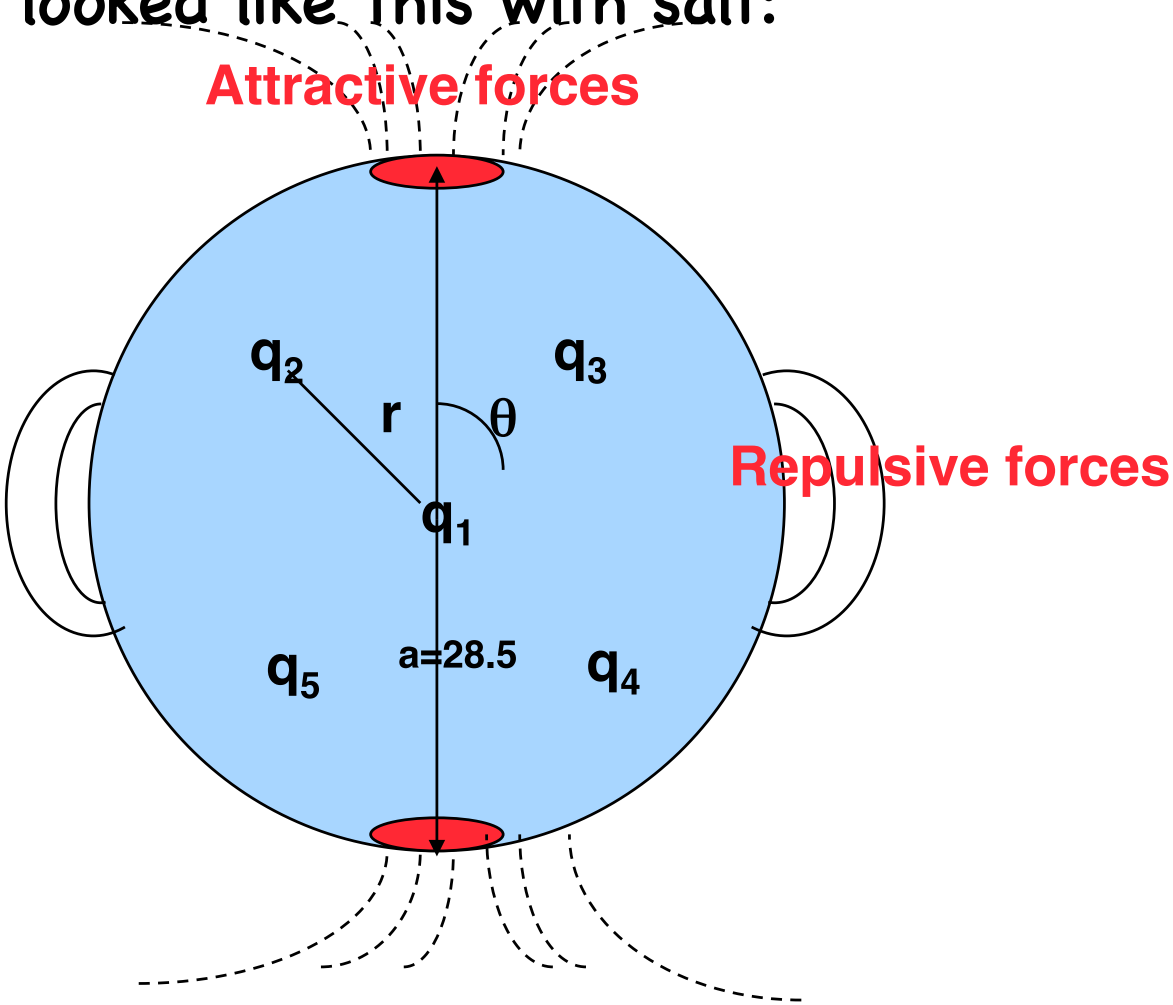
Summary of Results for SOD/O_2^-

The electrostatic field lines looked like this with salt:



Summary of Results for SOD/O_2^-

The electrostatic field lines looked like this with salt:

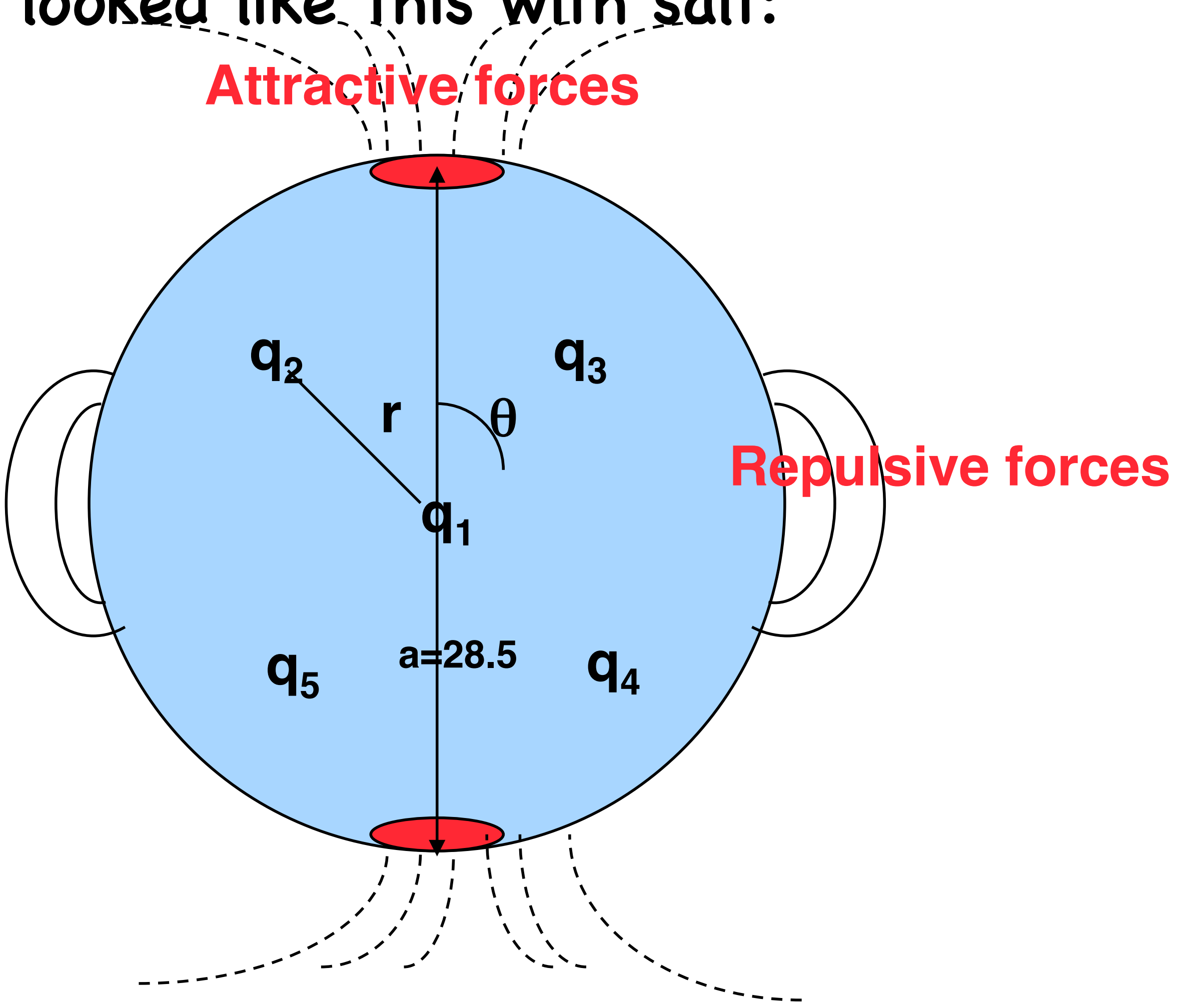


(3) The rate decreases with increasing salt concentration

The steering effect by SOD to guide superoxide toward protein active site is screened out. Salt decreases rate!

Summary of Results for SOD/O_2^-

The electrostatic field lines looked like this with salt:



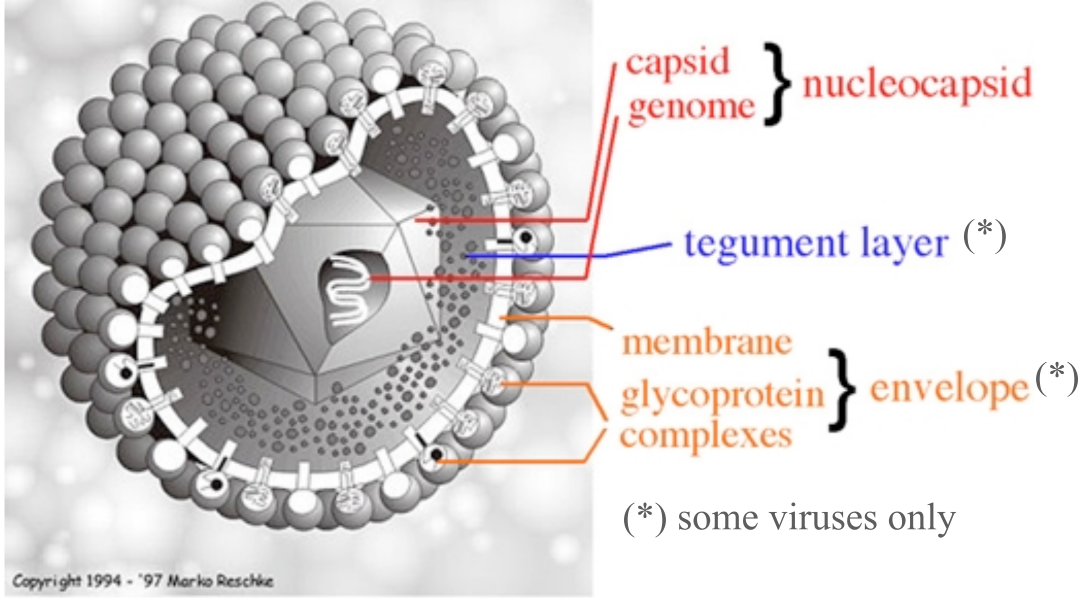
(3) The rate decreases with increasing salt concentration

The steering effect by SOD to guide superoxide toward protein active site is screened out. Salt decreases rate!

Mutations that destroy this charge distribution have the potential to destroy this steering effect

Virus Structure and Assembly

PRV virus, Dr. Marko Reschke, 1994



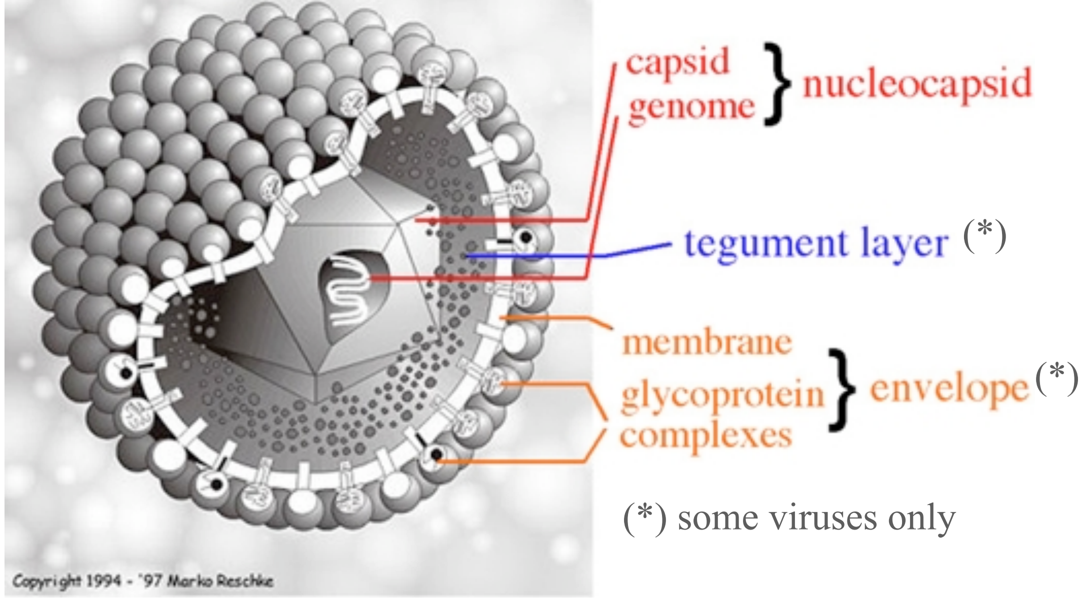
(*) some viruses only

Anti-viral drugs and vaccines and predicting new strains of viruses that arise from mutation are of tremendous clinical and fundamental importance

Simulating their assembly from their constituent proteins is of importance for designing better gene delivery systems

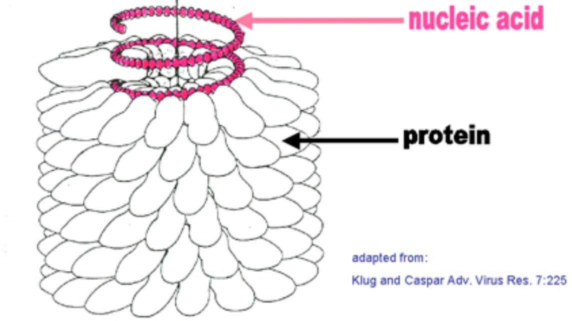
Virus Structure and Assembly

PRV virus, Dr. Marko Reschke, 1994



(*) some viruses only

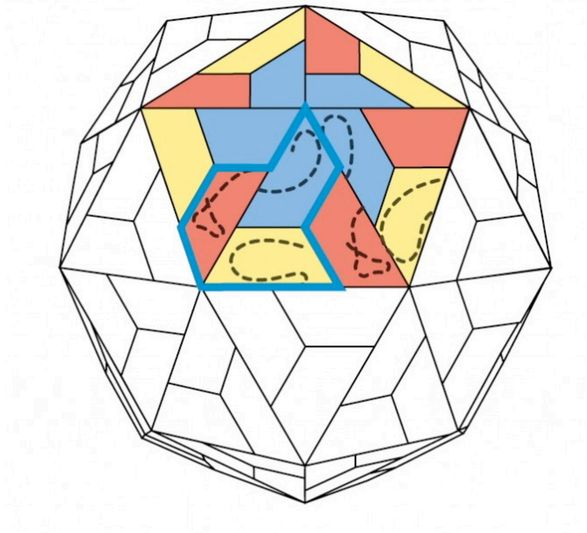
Helical capsid



adapted from:
Klug and Caspar Adv. Virus Res. 7:225

Icosahedral capsid

- 20 facets



Each facet:

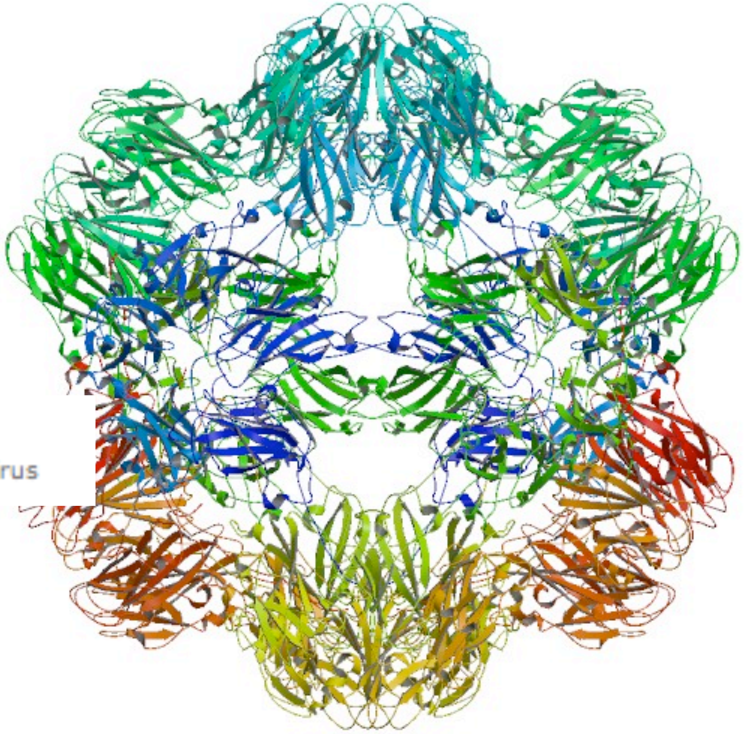
- *3-fold symmetry*
- *comprised of 3T monomers*

T = 'Triangulation number'

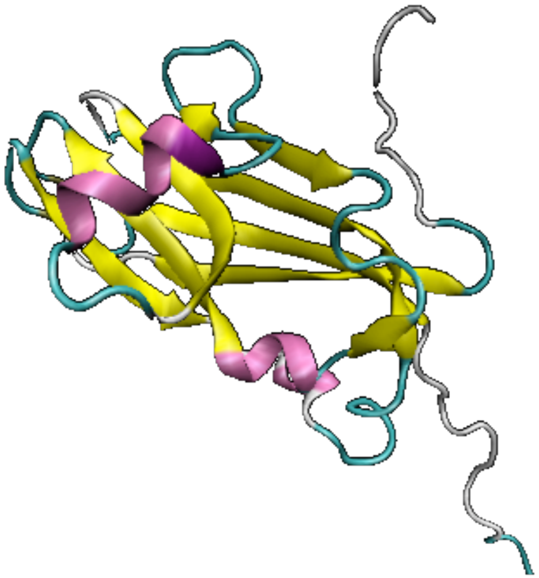
- *T=1 : 20x3x1 = 60 monomers*
- *T=3 : 20x3x3 = 180 monomers*

Brome Mosaic Virus

- plant virus: infects grass
- forms empty capsid *in vitro*
- potential drug delivery vehicle



Biological Assembly Image for 1YC6
Crystallographic Structure of the T=1 Particle of Brome Mosaic Virus

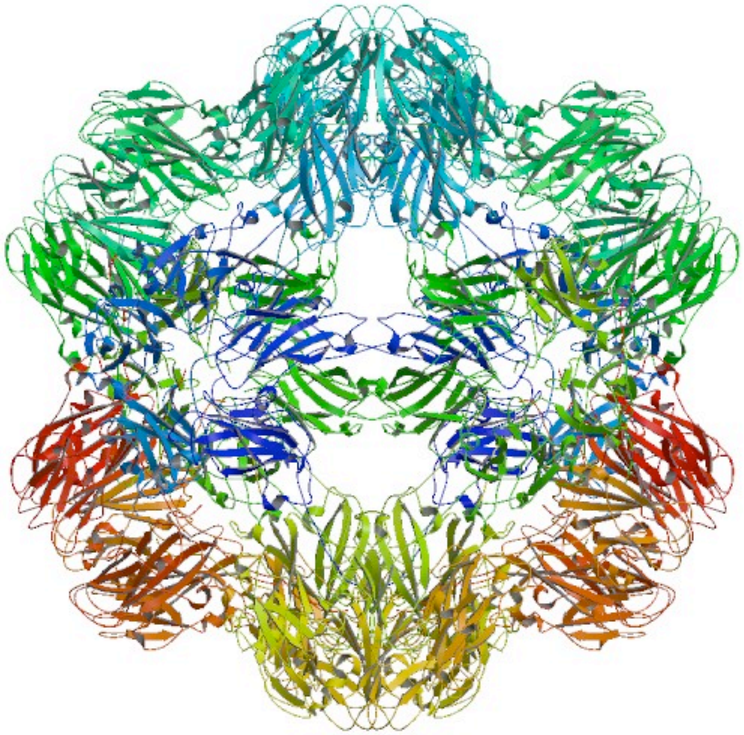


WT forms T=3 particle (189 residues/ monomer and 180 monomers

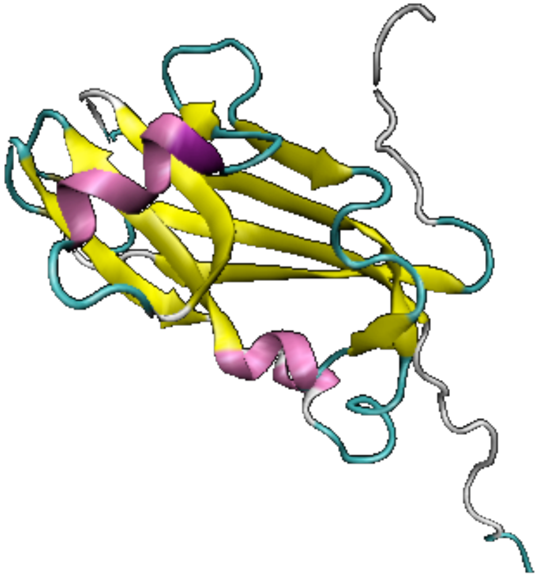
Mutant (1YC6) forms T=1 particle (154 residues/monomer and 60 monomers)

Brome Mosaic Virus

- plant virus: infects grass
- forms empty capsid *in vitro*
- potential drug delivery vehicle



Can we assemble the Brome Mosaic virus from its monomers in these two forms?



WT forms T=3 particle (189 residues/ monomer and 180 monomers

Mutant (1YC6) forms T=1 particle (154 residues/monomer and 60 monomers)

Solving the Linearized PB Equation

- Analytical (Multipole) Methods
  - Debye-Huckel; Kirkwood (1934): 1 molecule

Solving the Linearized PB Equation

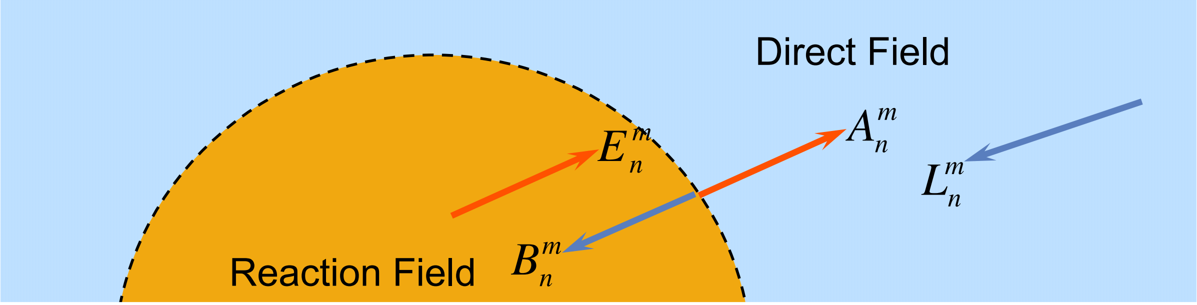
- Analytical (Multipole) Methods
    - Kirkwood (1934): 1 molecule
    - Lotan & THG (2006): N molecules! Generalized theory of Kirkwood (1934) to arbitrary numbers of molecules through advances in multipole methods for Yukawa potential (Greengard and co-workers)
    - Fast and accurate
    - Restricted to idealized geometries

Solving the Linearized PB Equation

- Numerical Methods (Arbitrary dielectric geometries)
  - Finite Difference Methods (FD): Delphi, UHBD, APBS
  - Finite Element Methods (FE): APBS
    - Realistic geometries; sparse matrix solvers available
    - More recent (APBS) are very fast
    - Mesh does not follow dielectric boundary exactly
    - does not include von Neumann boundary condition
    - forces estimated from finite difference of potential

  - Boundary Element Methods (BEM)
    - Surface (2D) instead of volume (3D): less unknowns
    - both Dirchlet and Von Neumann b.c. satisfied by construction
    - Memory Intensive; computationally intensive, can handle only few macromolecules

Analytical Solution to PBE for N Molecules



Multipole expansion of charge:

$$E_{n,m}^{(i)} = \sum_{j=1}^{M_i} q_j^{(i)} \left(\rho_j^{(i)}\right)^n Y_{n,m}\left(\vartheta_j^{(i)}, \varphi_j^{(i)}\right)$$

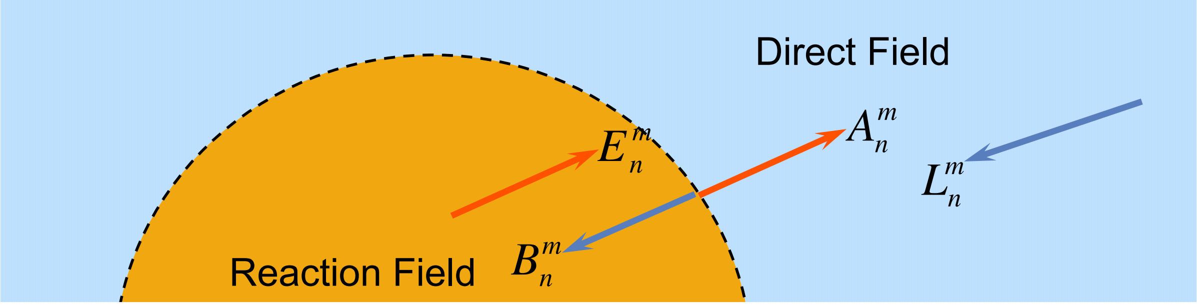
Potential inside cavity:

$$\Phi_{in}^{(i)}(\mathbf{t}) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \left(\frac{E_{n,m}^{(i)}}{\varepsilon_p r^{n+1}} + B_{n,m}^{(i)} r^n \right) Y_{n,m}(\theta,\phi)$$

Potential outside Cavity:

$$\Phi_{out}^{(i)}(\mathbf{t}) = \frac{1}{\varepsilon_s} \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \left(\frac{A_{n,m}^{(i)}}{r^{n+1}} e^{-\kappa r} \hat{k}_n(\kappa r) + L_{n,m}^{(i)} r^n \hat{i}_n(\kappa r) \right) Y_{n,m}(\theta,\phi)$$

Analytical Solution to PBE for N Molecules



Multipole expansion of charge: $E_{n,m}^{(i)} = \sum_{j=1}^{M_i} q_j^{(i)} \left(\rho_j^{(i)}\right)^n Y_{n,m}\left(\vartheta_j^{(i)}, \varphi_j^{(i)}\right)$

Potential inside cavity: $\Phi_{in}^{(i)}(\mathbf{t}) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \left(\frac{E_{n,m}^{(i)}}{\varepsilon_p r^{n+1}} + B_{n,m}^{(i)} r^n \right) Y_{n,m}(\theta,\phi)$

Potential outside Cavity: $\Phi_{out}^{(i)}(\mathbf{t}) = \frac{1}{\varepsilon_s} \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \left(\frac{A_{n,m}^{(i)}}{r^{n+1}} e^{-\kappa r} \hat{k}_n(\kappa r) + L_{n,m}^{(i)} r^n \hat{i}_n(\kappa r) \right) Y_{n,m}(\theta,\phi)$

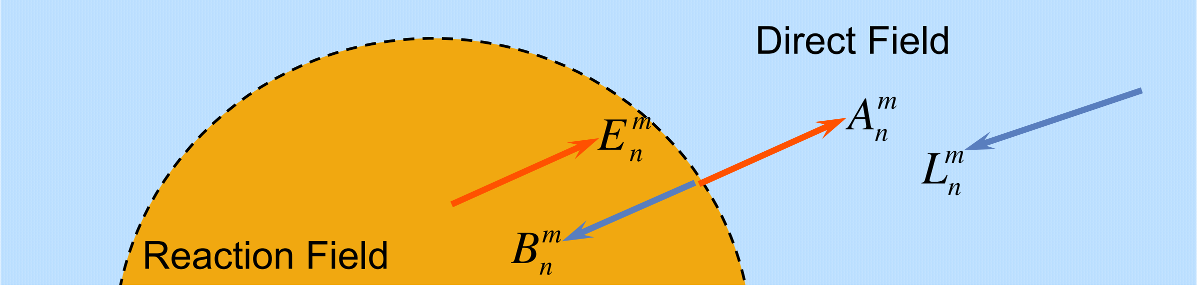
Linear re-expansion operator: $\mathbf{L}^{(i)} = \sum_{\substack{j=1 \\ j \neq i}}^{N} \mathrm{T}^{(i,j)} \cdot \mathbf{A}^{(j)}$

Multipole methods for Yukawa potential (Greengard et al)

Transforms a multipole expansion at $\mathbf{c}^{(j)}$ to a local (Taylor) expansion at $\mathbf{c}^{(i)}$.

Analytical Solution to PBE for N Molecules

Solving for **A**$^{(i)}$ and **B**$^{(i)}$:

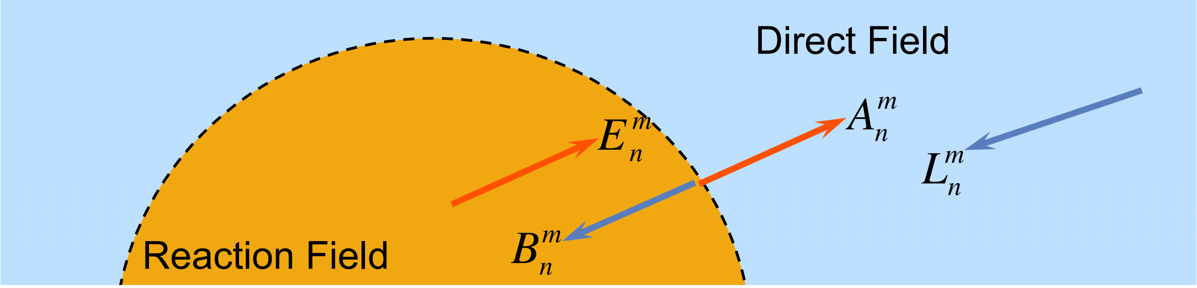


$$A_{n,m}^{(i)} = \gamma_n^{(i)}\delta_n^{(i)}L_{n,m}^{(i)} + \gamma_n^{(i)}E_{n,m}^{(i)}$$

$$B_{n,m}^{(i)} = \frac{1}{\varepsilon_s}\left(\frac{A_{n,m}^{(i)}}{a_i^{2n+1}}e^{-\kappa a_i}\hat{k}_n\left(\kappa a_i\right) + L_{n,m}^{(i)}\hat{i}_n\left(\kappa a_i\right)\right) - \frac{E_{n,m}^{(i)}}{\varepsilon_p a_i^{2n+1}}$$

If I solve **A**$^{(i)}$, then **B**$^{(i)}$ is straightforward; focus on solving **A**$^{(i)}$

Analytical Solution to PBE for N Molecules



Solving for **A**$^{(i)}$ and **B**$^{(i)}$:

$$A_{n,m}^{(i)} = \gamma_n^{(i)}\delta_n^{(i)}L_{n,m}^{(i)} + \gamma_n^{(i)}E_{n,m}^{(i)}$$

$$B_{n,m}^{(i)} = \frac{1}{\varepsilon_s}\left(\frac{A_{n,m}^{(i)}}{a_i^{2n+1}}e^{-\kappa a_i}\hat{k}_n(\kappa a_i) + L_{n,m}^{(i)}\hat{i}_n(\kappa a_i)\right) - \frac{E_{n,m}^{(i)}}{\varepsilon_p a_i^{2n+1}}$$

If I solve **A**$^{(i)}$, then **B**$^{(i)}$ is straightforward; focus on solving **A**$^{(i)}$

Define matrix $\mathbf{\Gamma}^{(i)}$

$$\gamma_n^{(i)} = \frac{(2n+1)e^{\kappa a_i}}{(2n+1)\hat{k}_{n+1}(\kappa a_i) + n\hat{k}_n(\kappa a_i)\left(\varepsilon_p/\varepsilon_s - 1\right)}$$

and matrix $\mathbf{\Delta}^{(i)}$

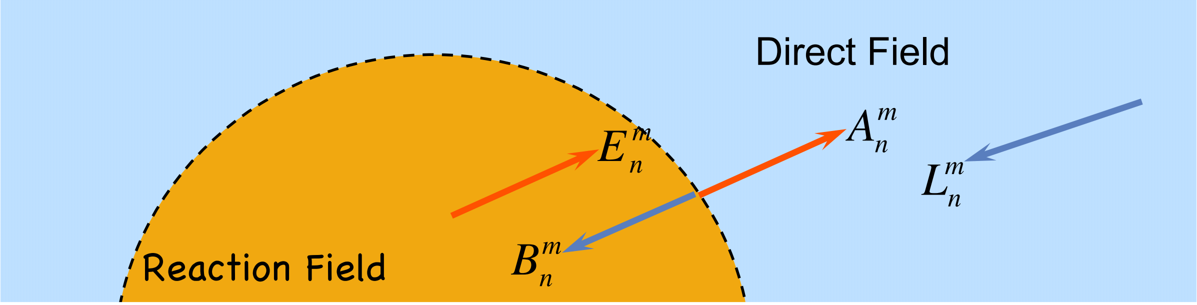
$$\delta_n^{(i)} = \frac{a_i^{2n+1}}{2n+1}\left[\kappa^2 a_i^2\frac{\hat{i}_{n+1}(\kappa a_i)}{2n+3} + n\hat{i}_n(\kappa a_i)\left(1 - \varepsilon_p/\varepsilon_s\right)\right]$$

Note that these matrices are merely constants of the system we are solving (and we solve once in a CG scheme)

Analytical Solution to PBE for N Molecules

Solution to $A^{(i)}$ can be written as

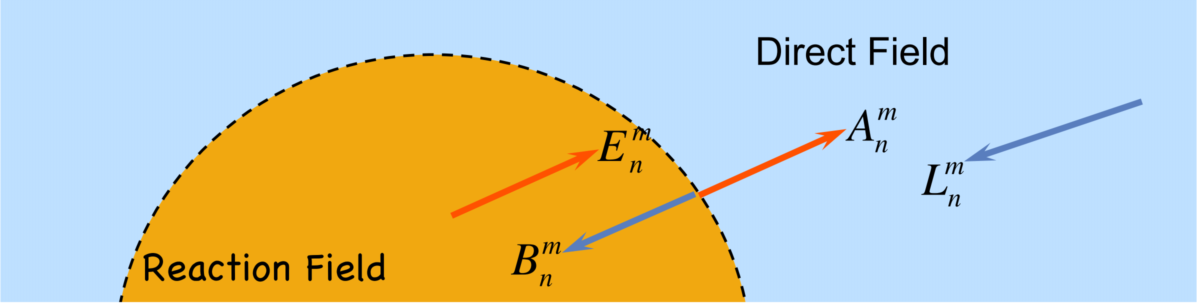
$$A = \Gamma\left(\Delta \cdot T \cdot A + E\right)$$



Analytical Solution to PBE for N Molecules

Solution to $A^{(i)}$ can be written as

$$A = \Gamma\left(\Delta \cdot T \cdot A + E\right)$$



$\Delta^{(i)}$ Cavity-polarization operator from external charges

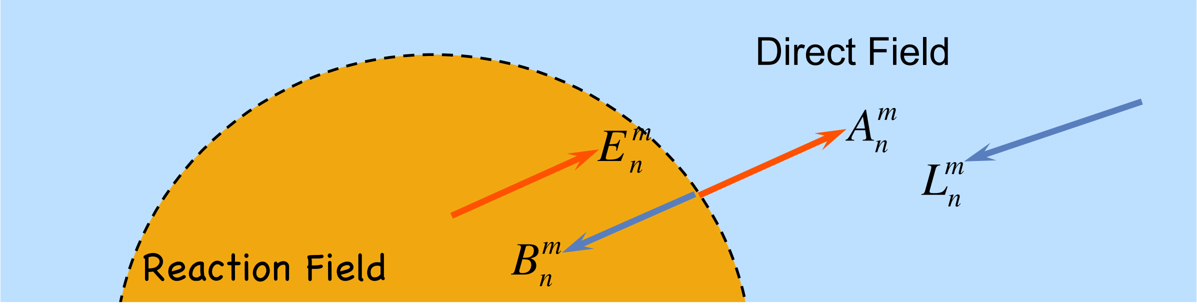
$\Gamma^{(i)}$ Dielectric-boundary-crossing operator for a molecule

$\mathbf{A}^{(i)}$ *Effective* multipole expansion: transformed charges due to dielectric boundary and polarization of other empty dielectric cavities

Analytical Solution to PBE for N Molecules

Solution to $A^{(i)}$ can be written as

$$A = \Gamma\left(\Delta \cdot T \cdot A + E\right)$$



$\Delta^{(i)}$ Cavity-polarization operator from external charges

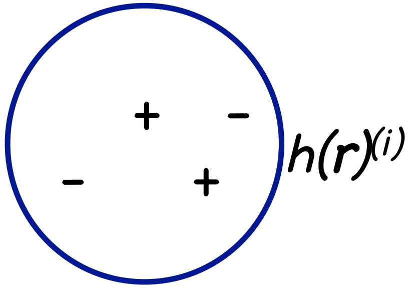
$\Gamma^{(i)}$ Dielectric-boundary-crossing operator for a molecule

$\mathbf{A}^{(i)}$ *Effective* multipole expansion: transformed charges due to dielectric boundary and polarization of other empty dielectric cavities

The potential field induced by a molecule is due to the sum of contribution of its free charges ($E^{(i)}$) and the contribution of polarization charges induced by other molecules, transformed by the effect of its dielectric boundary.

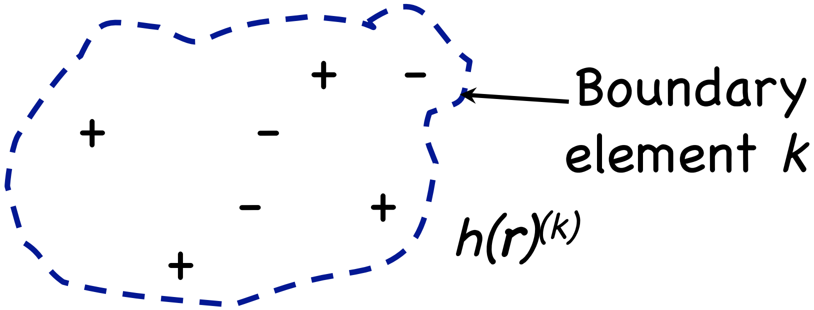
Semi-analytical PBE: N Realistic Shapes

Multipole Method



Solve for entire molecule *i* analytically

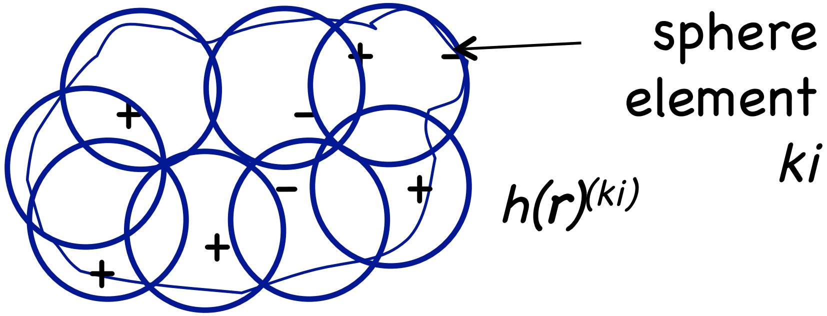
Boundary Element Method



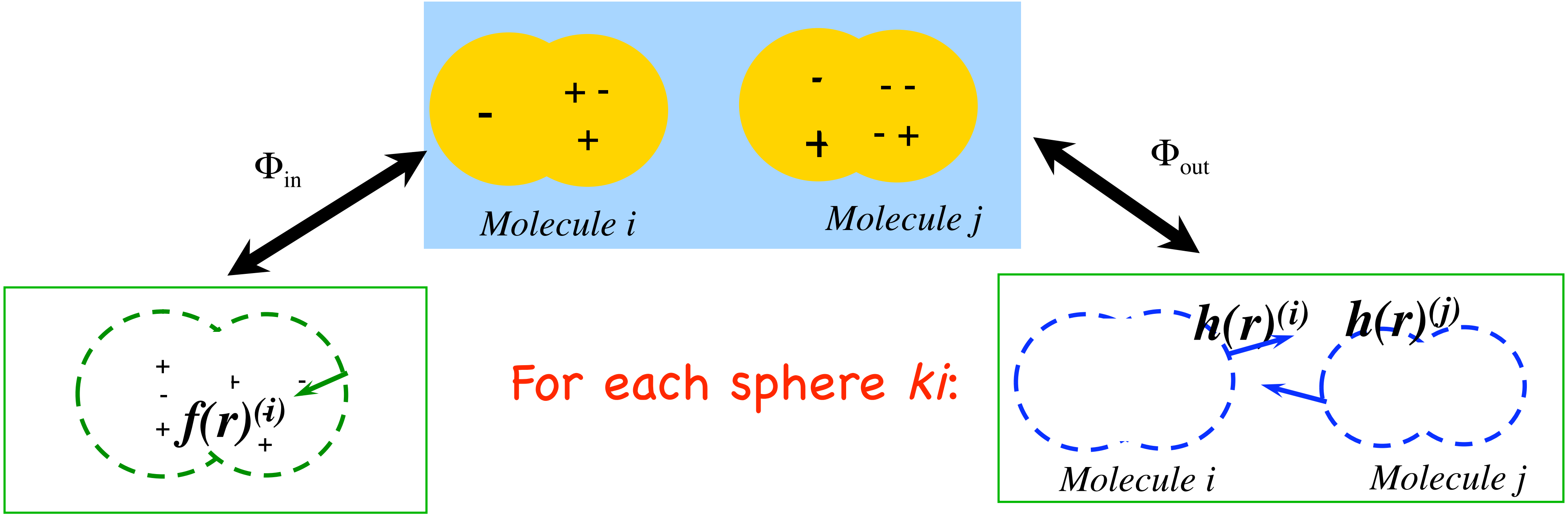
Solve for N_k system of equations numerically $N_k \sim 10^6$

Multipole Method for arbitrary shapes

Discretize molecule into N_{ki} spheres
Solve for each sphere *ki* analytically,
Iterate over N_{ki} spheres, $N_{ki} \sim 100$



Semi-analytical PBE: N Realistic Shapes



For each sphere *ki*:

Express Φ_{in} and Φ_{out} in terms of multipoles $\mathbf{F}^{(ki)}$ and $\mathbf{H}^{(ki)}$

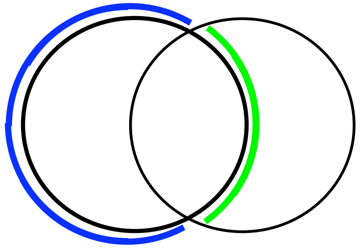
Impose boundary equations and solve for $\mathbf{F}^{(ki)}$ and $\mathbf{H}^{(ki)}$

Update contribution from $\mathbf{F}^{(ki)}$ and $\mathbf{H}^{(ki)}$ to other spheres

Repeat for all spheres until convergence criteria is reached

Semi-analytical PBE: N Realistic Shapes

On the *EXPOSED* sphere surface (a,θ_E,ϕ_E):



$$\sum_{n=0}^{\infty}\sum_{m=-n}^{n}\left(-H_{nm}^{(ki)}e^{-\kappa a_{ki}}\hat{k}_n(\kappa a_{ki})+F_{nm}^{(ki)}+X_{Hnm}^{(ki)}\right)Y_{nm}(\theta_E,\phi_E)=0$$

$$\sum_{n=0}^{\infty}\sum_{m=-n}^{n}\left(e^{-\kappa a_{ki}}\left[n\hat{k}_n(\kappa a_{ki})-(2n+1)\hat{k}_{n+1}(\kappa a_{ki})\right]H_{nm}^{(ki)}-n\varepsilon F_{nm}^{(ki)}+X_{Fnm}^{(ki)}\right)Y_{nm}(\theta_E,\phi_E)=0$$

On the *BURIED* sphere surface (a,θ_B,ϕ_B):

$$\sum_{n=0}^{\infty}\sum_{m=-n}^{n}\frac{(2n+1)}{4\pi}F_{nm}Y_{nm}(\theta_B,\phi_B)=0$$

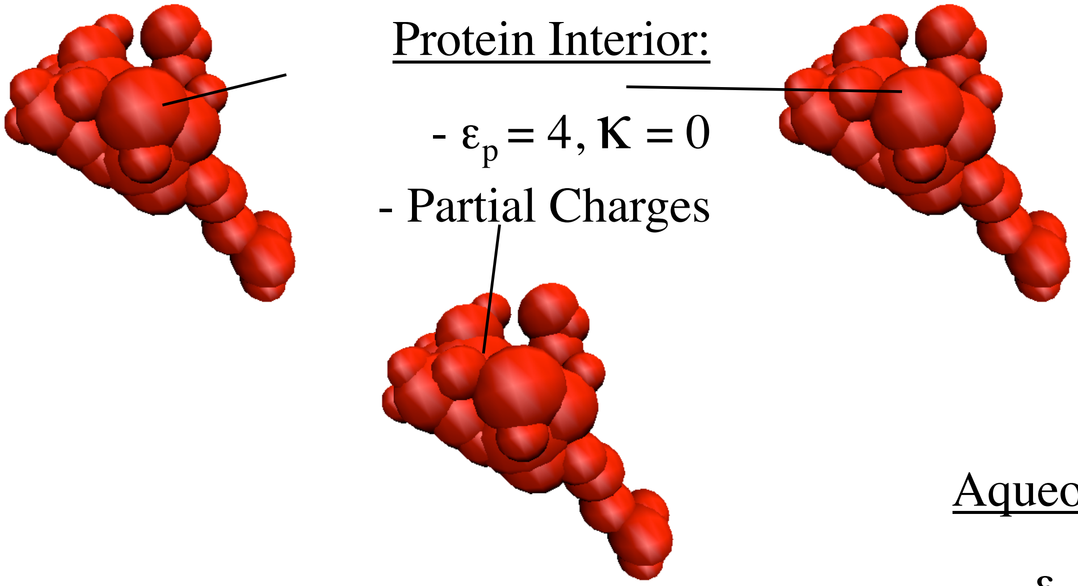
$$\sum_{n=0}^{\infty}\sum_{m=-n}^{n}\frac{2n+1}{4\pi}\frac{H_{nm}}{\hat{i}_n(\kappa a_{ki})}Y_{nm}(\theta_B,\phi_B)=0$$

Now we can solve for $F^{(ki)}$ and $H^{(ki)}$

Semi-analytical PBE: N Realistic Shapes

Each *protein* modeled as a cluster of overlapping Spheres

Each *sphere* contains fixed partial charges



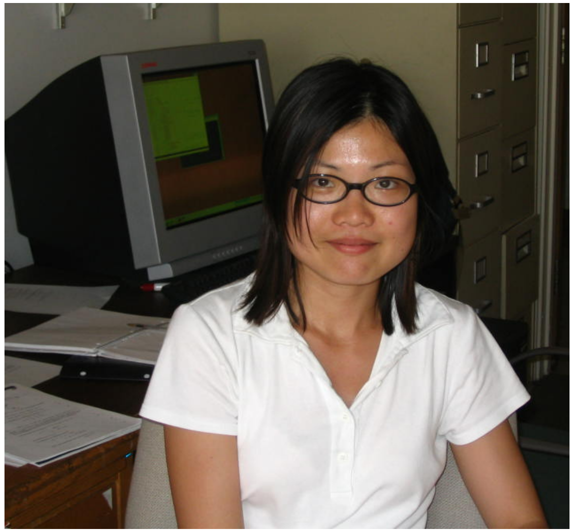
Protein Interior:

- $\varepsilon_p = 4, \kappa = 0$
- Partial Charges

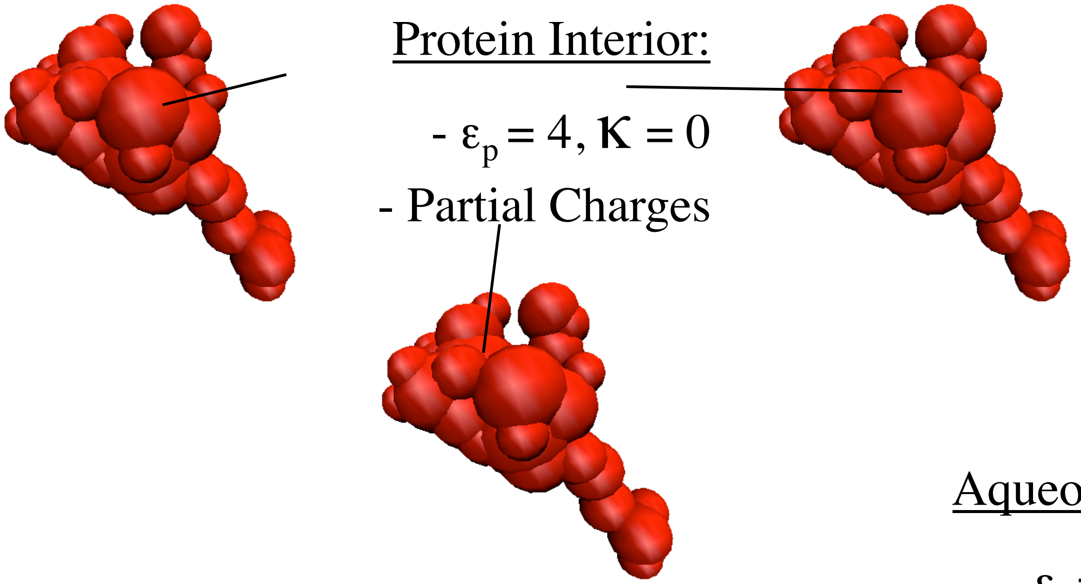
Aqueous solvent:

$\varepsilon_s = 78, \kappa > 0$

We are currently investigating the assembly of the capsid



Semi-analytical PBE: N Realistic Shapes



Protein Interior:

- $\varepsilon_p = 4, \kappa = 0$

- Partial Charges

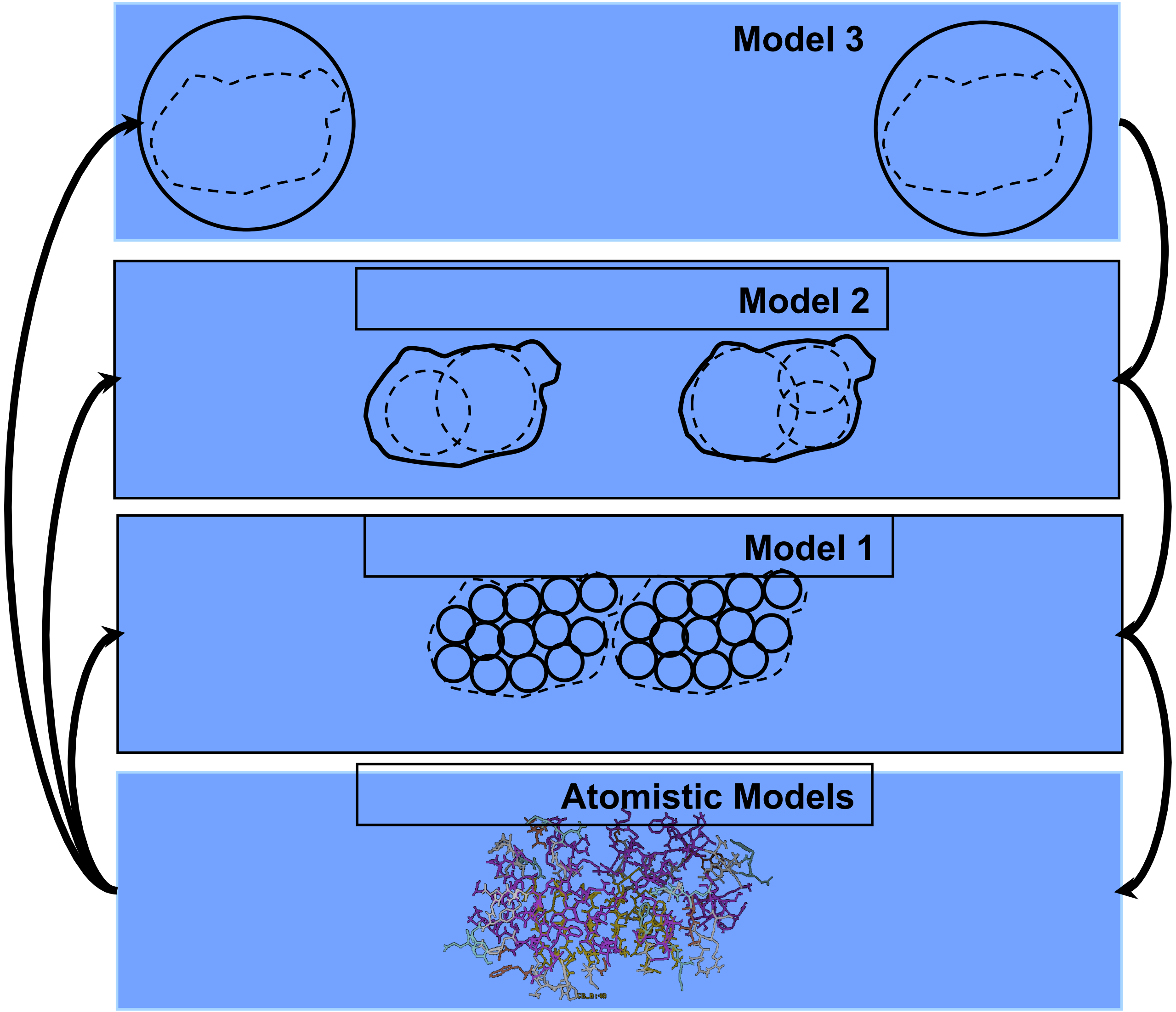
Aqueous solvent:

$\varepsilon_s = 78, \kappa > 0$

We are currently investigating the assembly of the capsid

Mulltiscale models

Proposed transitions between models



Spherical dielectric approximation breaks down

Separation d ~ (R_A , R_B)

Separation d ~ few water molecules