BIOL 555: Simulations of Macromolecular Structure

Sameer Varma
Research Assistant Professor
Department of Biological, Chemical and Physical Sciences
Illinois Institute of Technology, Chicago

Everything that living things do can be understood in terms of the jigglings and wigglings of atoms - Richard Feynman.

Role of molecular simulations

Combining molecular-simulation techniques with experiment accelerates fundamental as well as applied research, as molecular simulations

(a) provide mechanistic insights that complement experiment, and
(b) serve as an inexpensive route to scan the role of multiple physical parameters and environmental variables.

Computational techniques have become an integral part of biological research and discovery:

(a) Predicting bio-molecular (protein/RNA/lipid) structure
(b) Refining X-ray, NMR and neutron diffraction data
(c) Post-processing structural information
    - Interpret and analyze experimental data in terms of interactions at the atomic level
    - Understand fast time scale dynamics, such as during enzyme catalysis
(d) Structure-based drug design

Hierarchy of molecular simulation techniques

Quantum Monte Carlo

Post Hartree-Fock (MP2)

Kohn-Sham DFT

Quantum models

(picoseconds)

\+ Experimental data

calibrate

Explicit solvent

Implicit solvent

Atomistic models

(microseconds)

\+ Experimental data

Statistical data

calibrate

Mechanical models

(seconds)

calibrate

Coarse-grained / Mean-field models

(milliseconds)

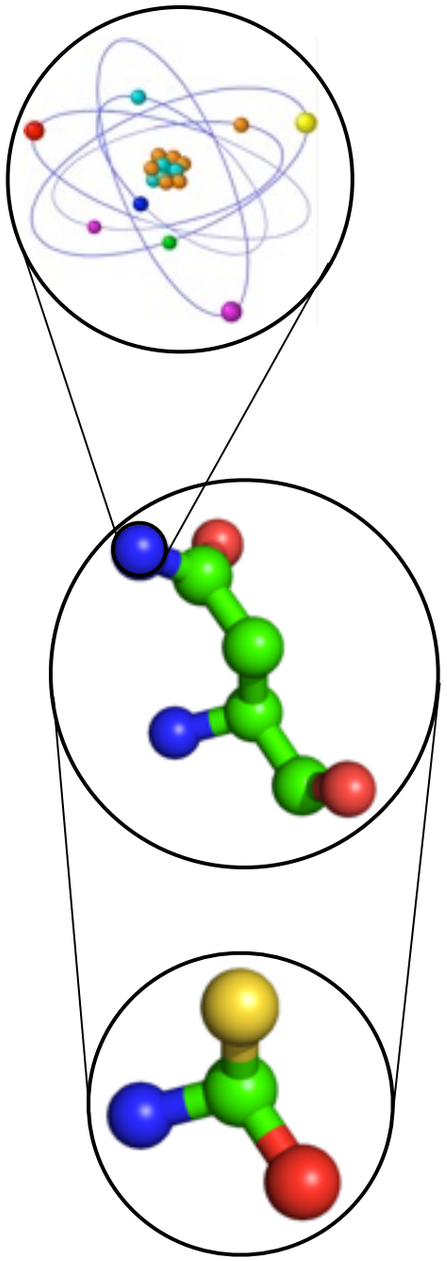
calibrate

Kinetic models

(hours)

Experimental data

Resolution



Access timescales: integrate classical equations of motion

Molecular Dynamics: Integrate Newton's/Hamilton's equation of motion

$$m_i \frac{\partial^2 \boldsymbol{r}_i}{\partial t^2} = \boldsymbol{F}_i,\ i = 1 \ldots N. \qquad \boldsymbol{F}_i = -\frac{\partial V}{\partial \boldsymbol{r}_i}$$

Langevin Dynamics: The 'unimportant' degrees of freedom are averaged-out in such a way that the thermodynamic and long time-scale properties are preserved.The reduction of degrees of freedom depends on the problem one wishes to solve. The interactions change into potentials of mean force, and the omitted degrees of freedom are replaced by noise and friction.

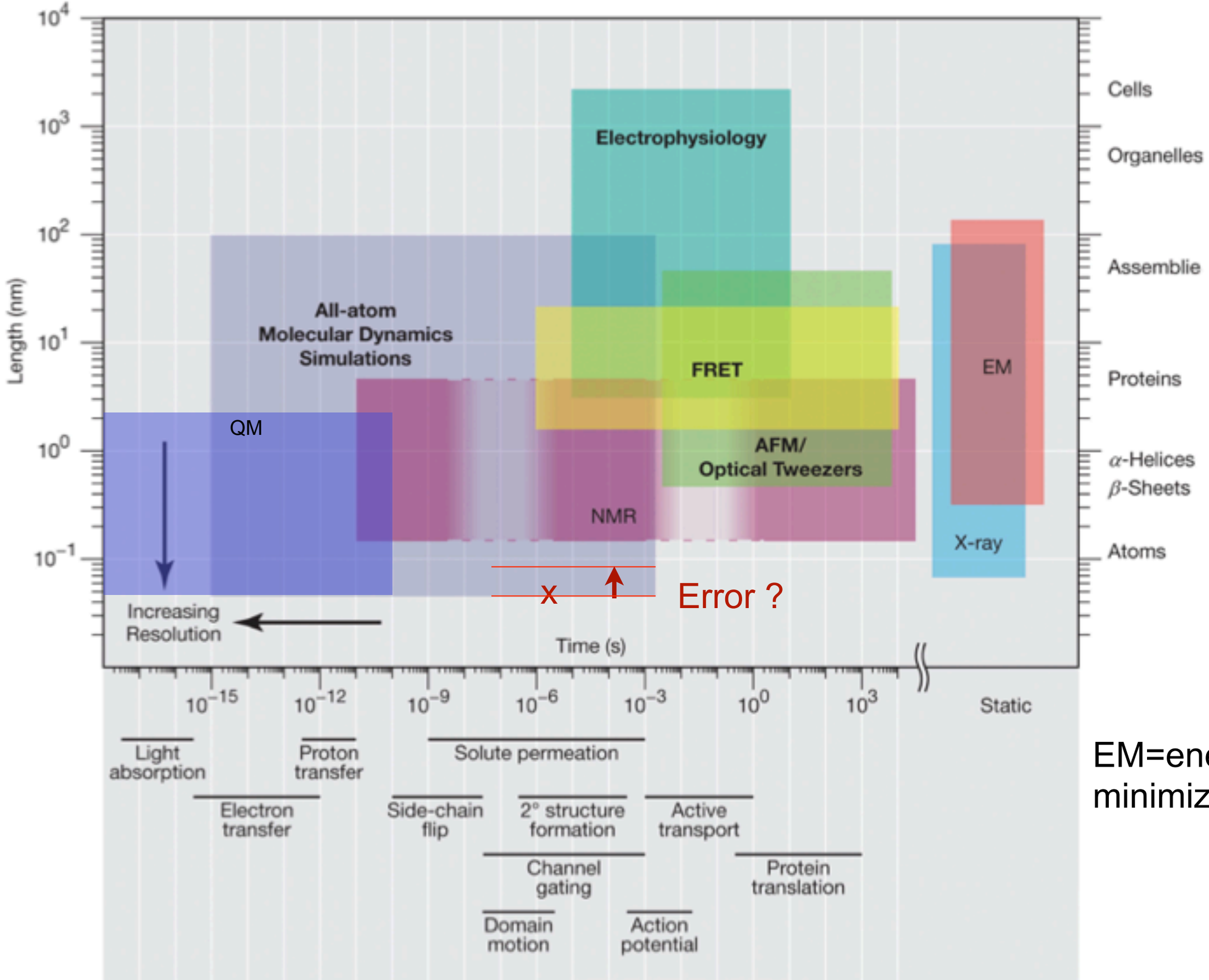
$$\boldsymbol{F}_i(t) = -\frac{\partial V^{\mathrm{mf}}}{\partial \boldsymbol{r}_i'} + \boldsymbol{F}_i^{\mathrm{friction}} + \boldsymbol{F}_i(t)^{\mathrm{noise}}$$

Brownian Dynamics: Langevin dynamics at high friction.

Monte Carlo: are a class of computational algorithms that rely on repeated random sampling to compute their results

Energy Minimization: No time step. Move atoms so as to reduce the net forces (the gradients of potential energy) on the atoms until they become negligible.

Comparing molecular simulations against experiment



Adopted from Shaw and Co. JGP 2010

EM=energy minimization

Define the potential energy V in atomistic simulations

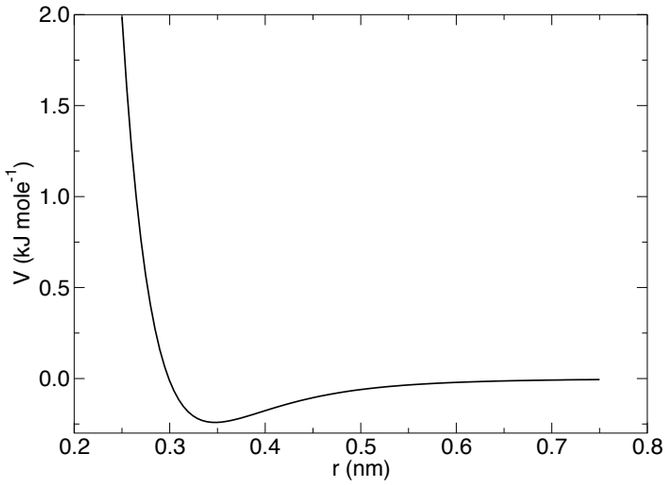
Pairwise additive assumption

$$V(\boldsymbol{r}_1, \ldots \boldsymbol{r}_N) = \sum_{i<j} V_{ij}(\boldsymbol{r}_{ij});$$

$$\boldsymbol{F}_i = -\sum_j \frac{dV_{ij}(r_{ij})}{dr_{ij}} \frac{\boldsymbol{r}_{ij}}{r_{ij}} = -\boldsymbol{F}_j$$

Non-bonded interactions: Eg. Lennard-Jones interaction

$$V_{LJ}(r_{ij}) = \frac{C_{ij}^{(12)}}{r_{ij}^{12}} - \frac{C_{ij}^{(6)}}{r_{ij}^{6}}$$



Coulomb interaction

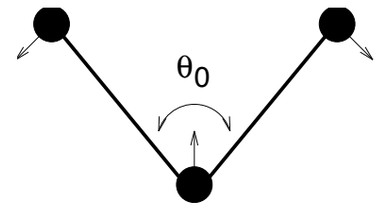
$$V_c(r_{ij}) = f\frac{q_i q_j}{\varepsilon_r r_{ij}}$$

Bond stretching: Eg. Harmonic potential

$$V_b\,(r_{ij}) = \frac{1}{2}k_{ij}^b(r_{ij} - b_{ij})^2$$

Define the potential energy V in atomistic simulations

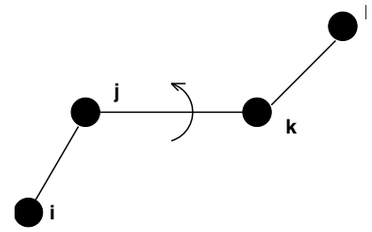
Angle stretching: Eg. Harmonic potential

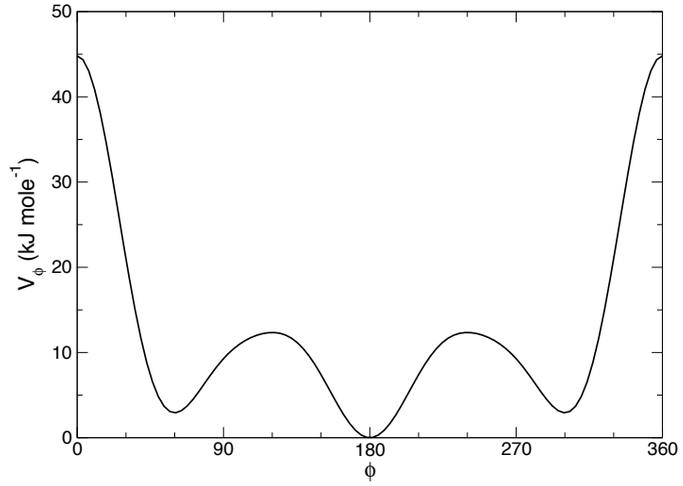


$$V_a(\theta_{ijk}) = \frac{1}{2} k^{\theta}_{ijk}(\theta_{ijk} - \theta^0_{ijk})^2$$

Dihedral potential: Eg. Ryckaert-Bellemans function

$$V_{rb}(\phi_{ijkl}) = \sum_{n=0}^{5} C_n(\cos(\psi))^n$$

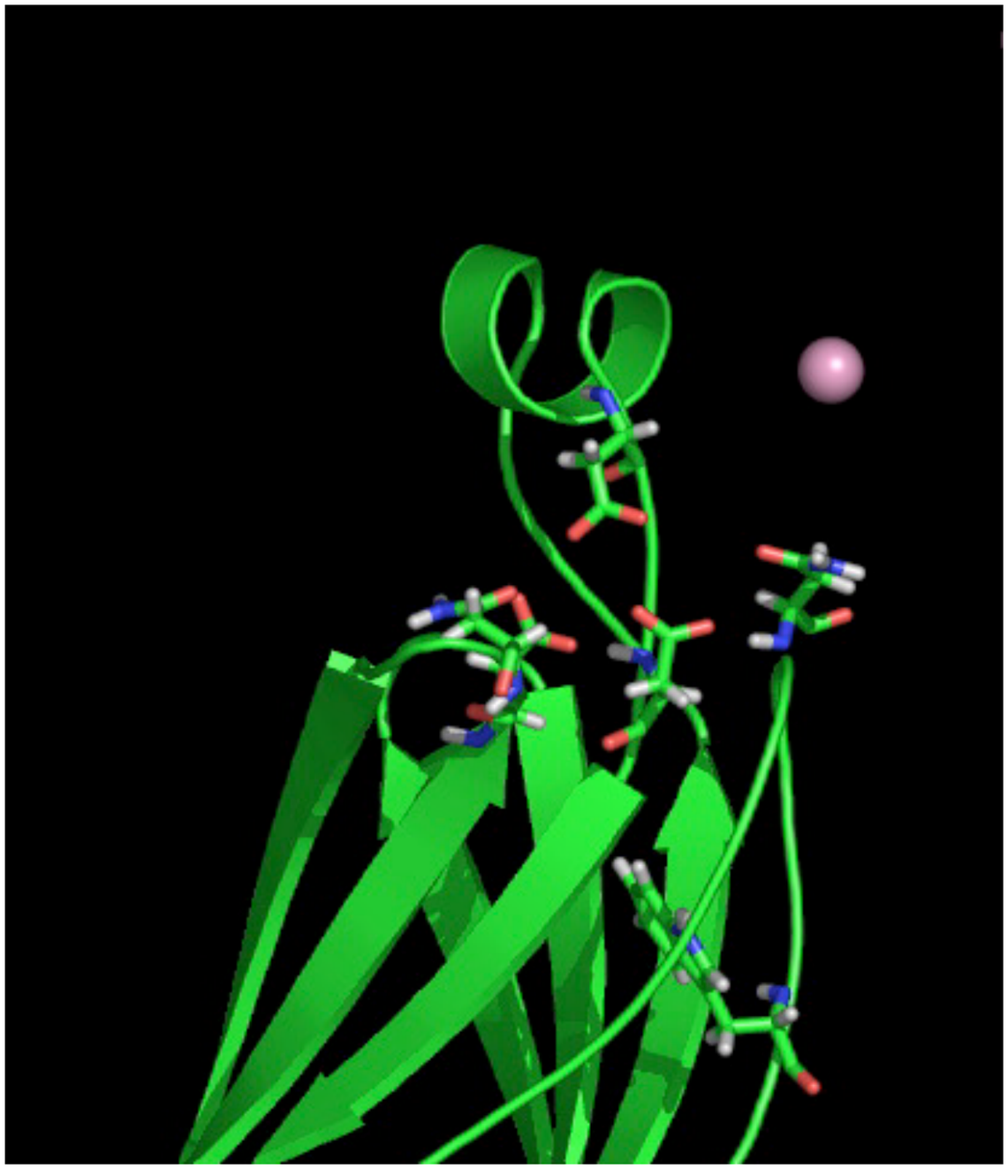




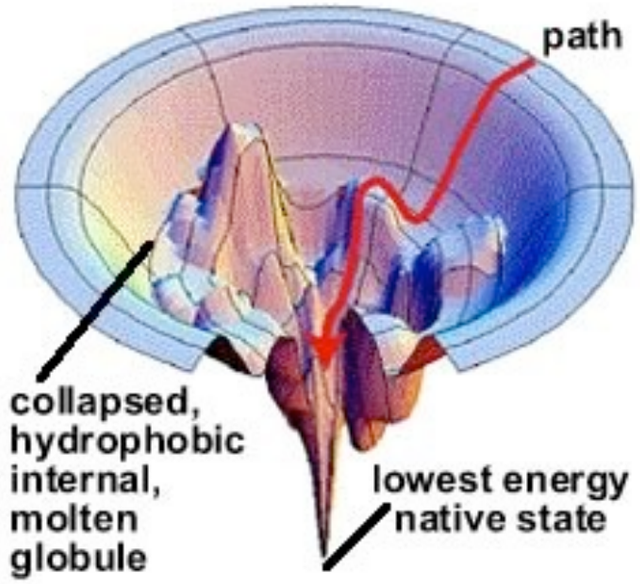
Polarization effects: Eg. Thole formulation

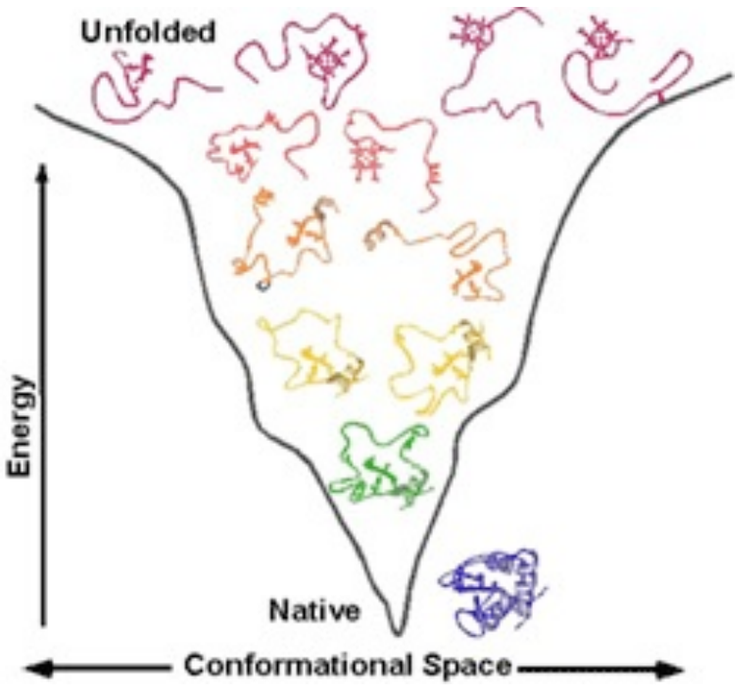
$$V_{thole} = \frac{q_i q_j}{r_{ij}} \left[1 - \left(1 + \frac{\bar{r}_{ij}}{2} \right) \exp^{-\bar{r}_{ij}} \right]$$

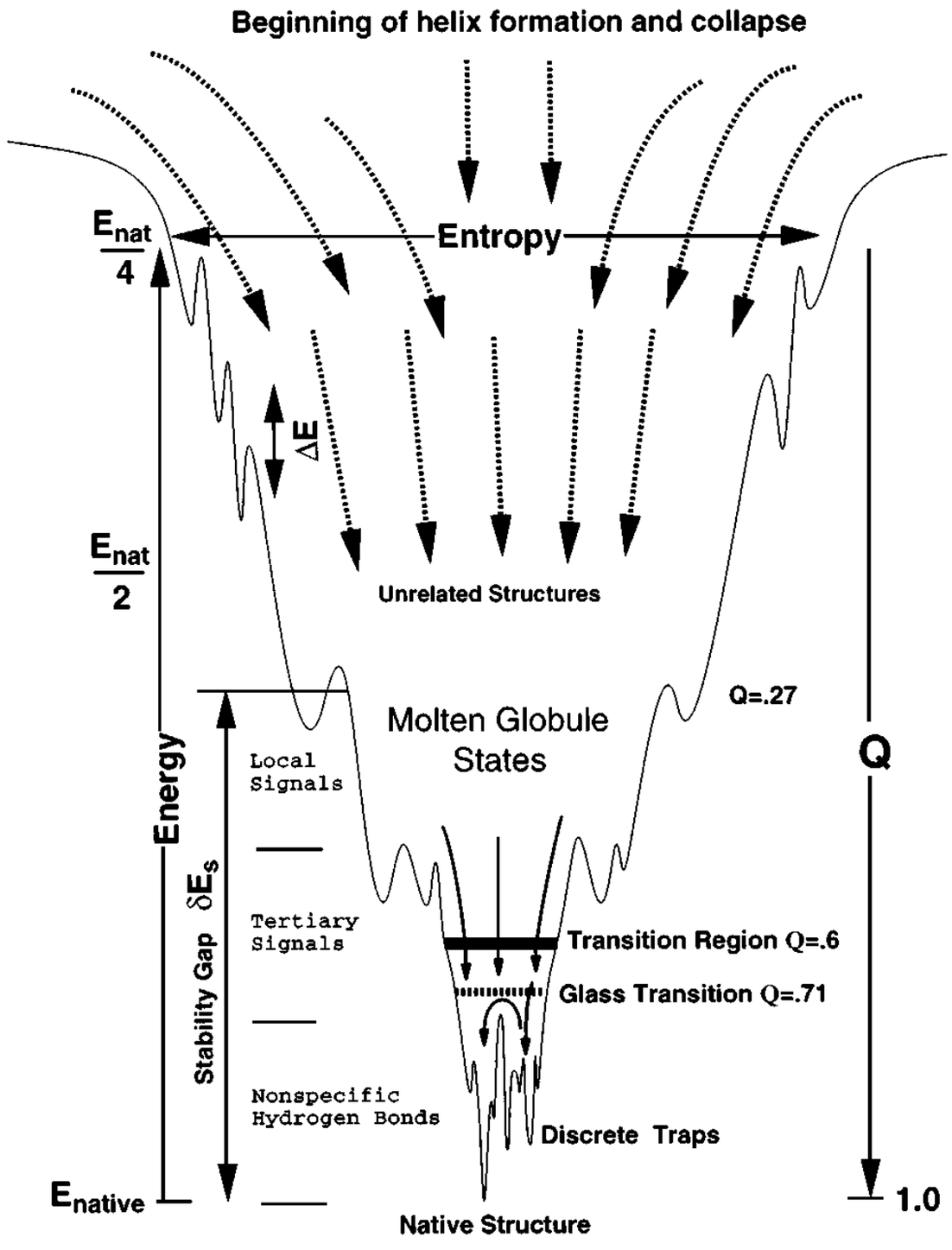
...and a whole bunch of restraints can be added to address a given problem.



Protein folding landscape







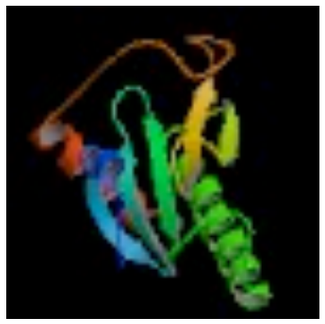
THEORY OF PROTEIN FOLDING *Annu. Rev. Phys. Chem. 1997. 48:545–600*

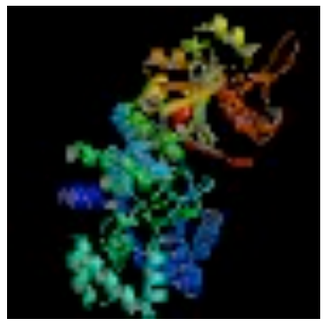
Critical Assessment of protein Structure Prediction (CASP)

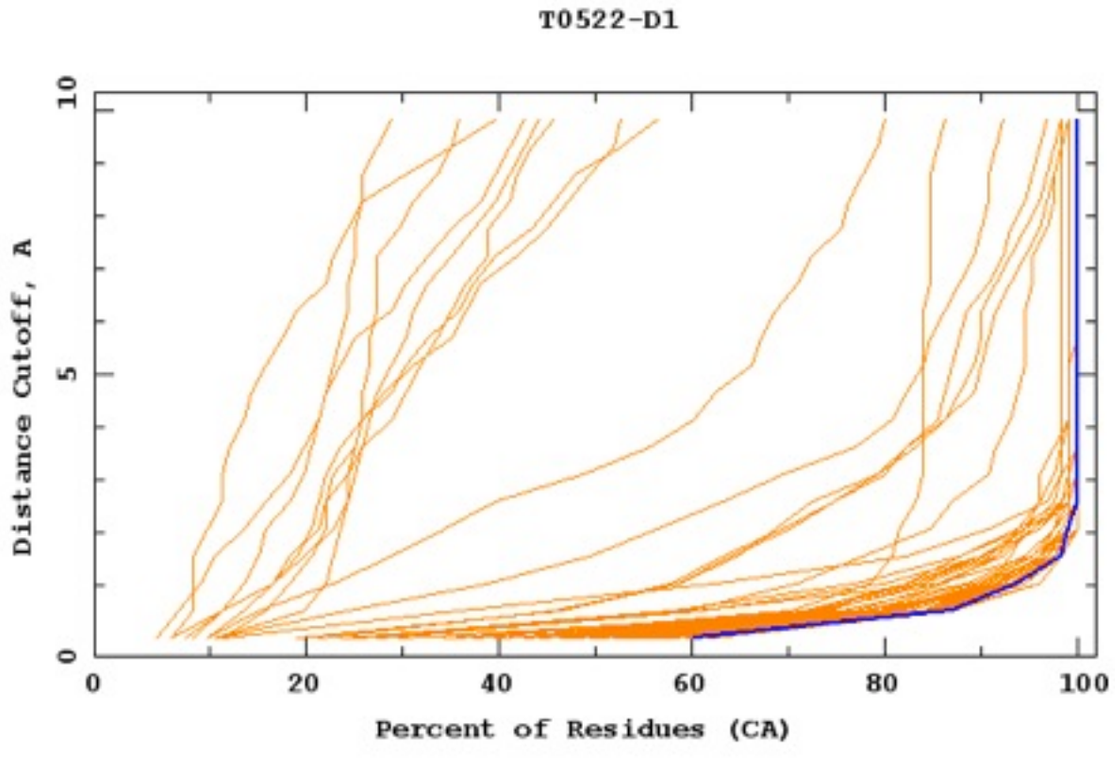
http://www.predictioncenter.org

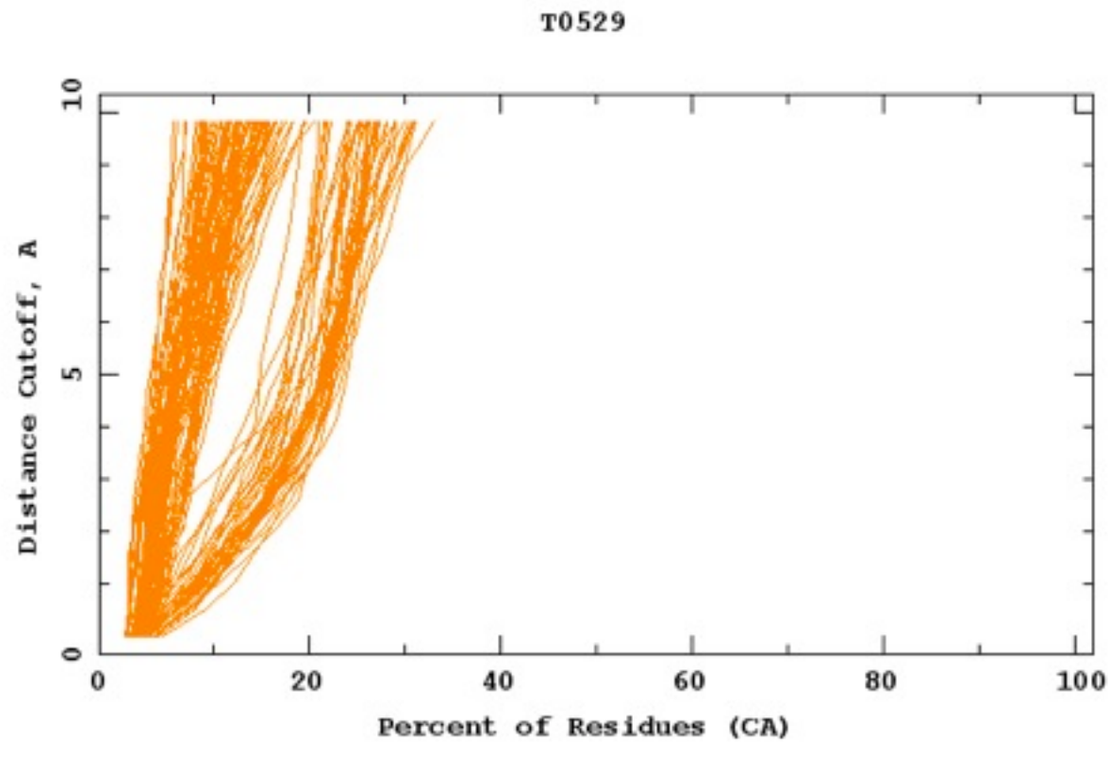
This protein folding challenge aims at establishing the current state of the art in protein structure prediction, identifying what progress has been made, and highlighting where future effort may be most productively focused.

CASP 9, 2010 results: 2 representative examples

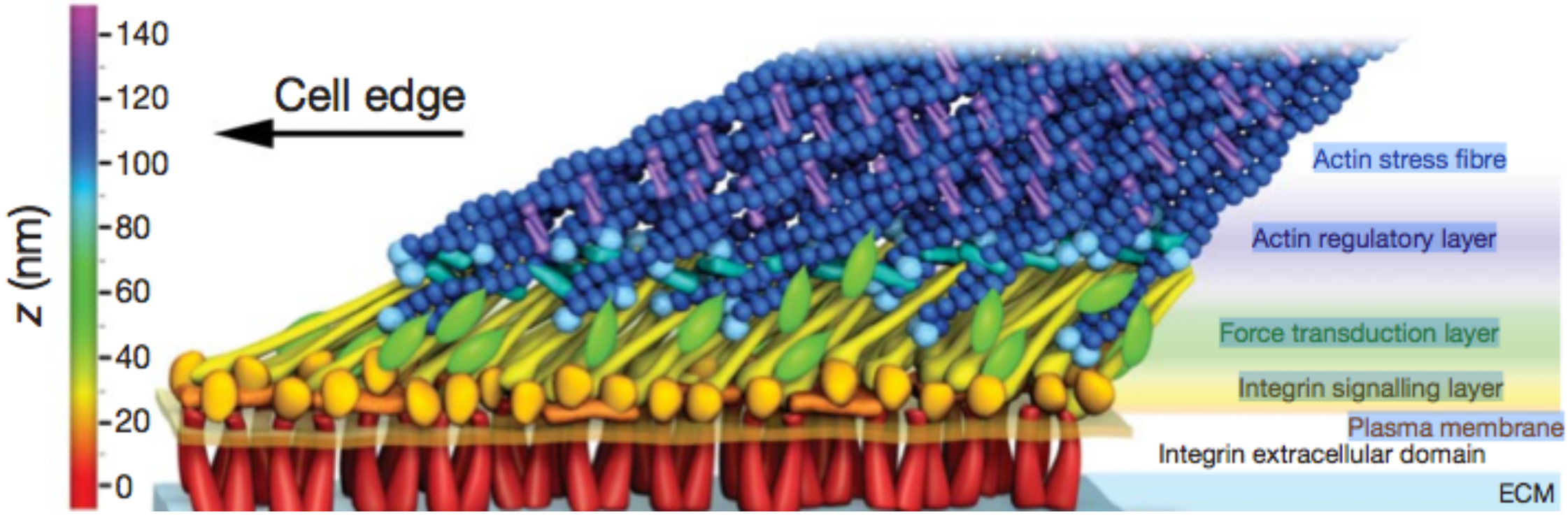




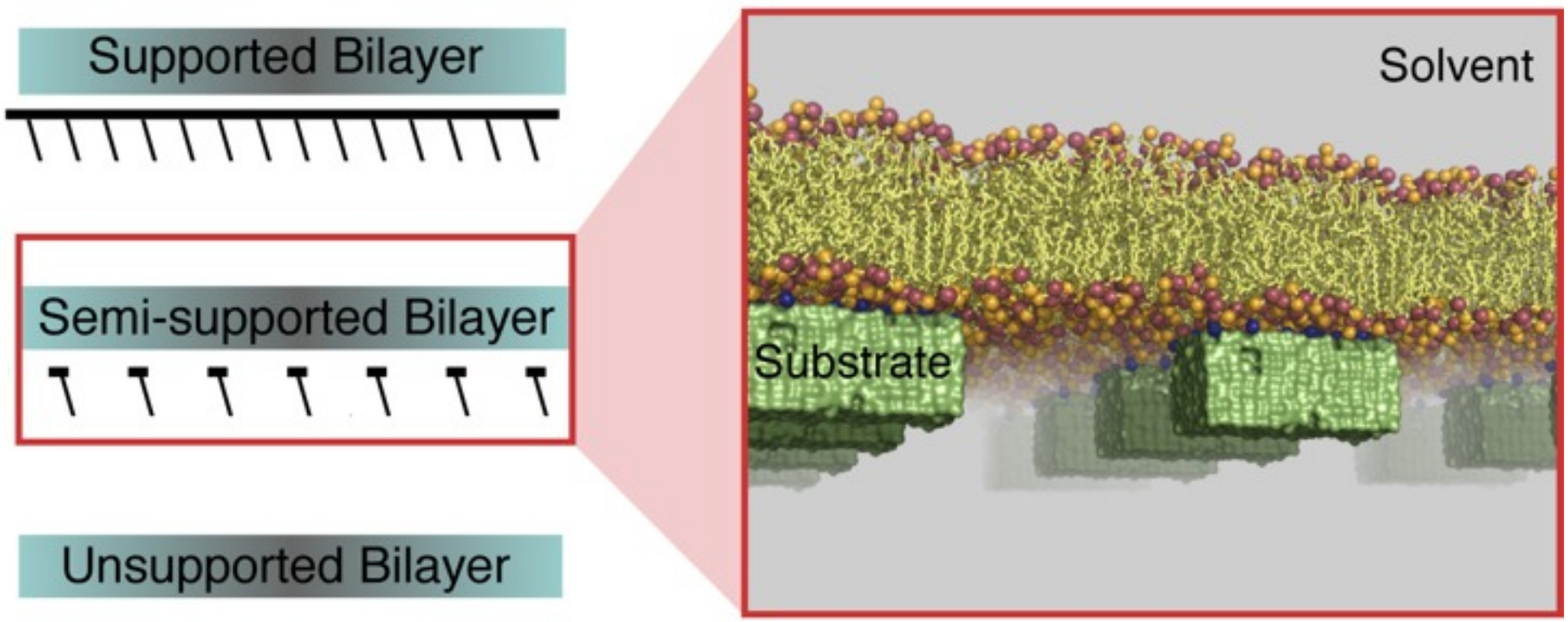




Membrane structure prediction - atomistic simulations



Kanchanawong *et al*. Nature 2010

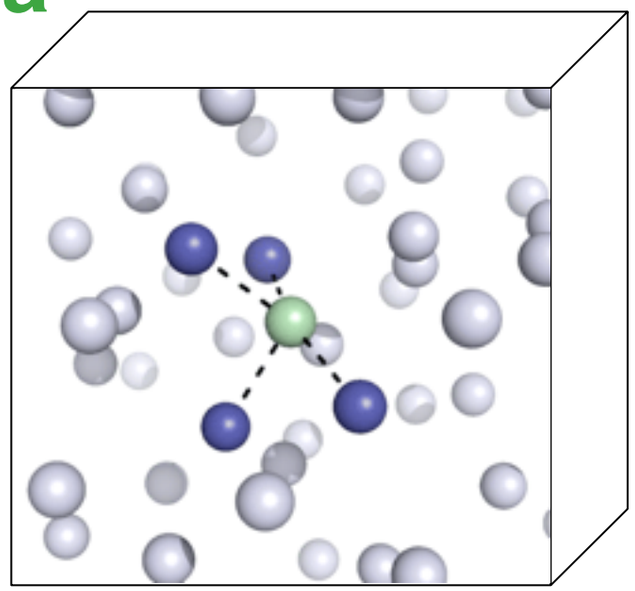


Varma and Scott, 2011

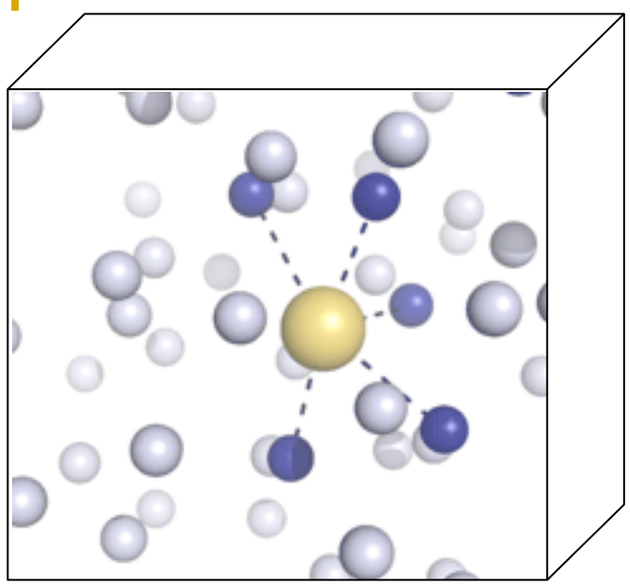
Structure/dynamics of ions in water: quantum mechanical simulations

Quantum chemical simulations of Ions in **Water**

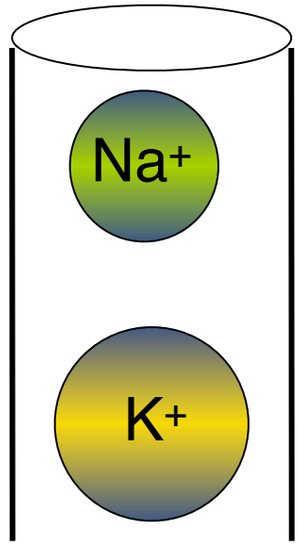
Na^+



K^+



	Pauling radii	Average ion-water distance in aqueous phase
Na^+	0.99 Å	2.4 Å
K^+	1.33 Å	2.8 Å



AIMD trajectory generated using

PW91 functional implemented in VASP 4.2,
NVE ensemble, PAW method, PME with background charge, Cut-off$_{KE}$= 500 eV

Varma and Rempe, 2006

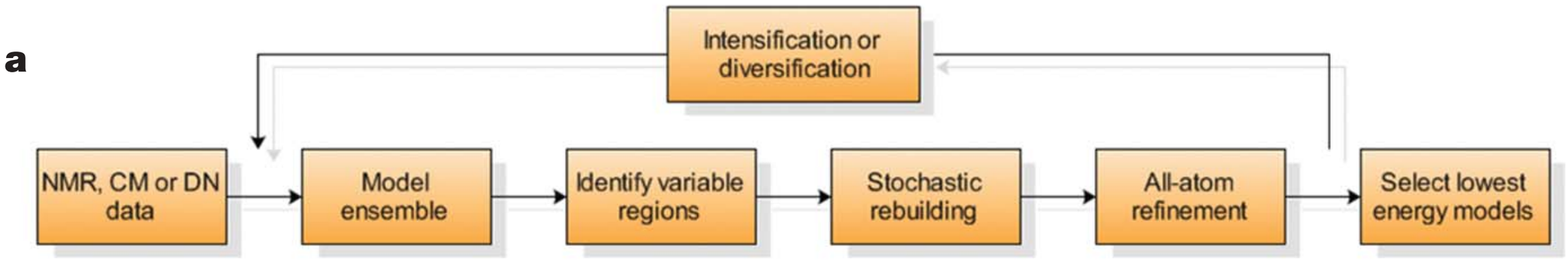
Vol 450|8 November 2007|doi:10.1038/nature06249 nature

ARTICLES

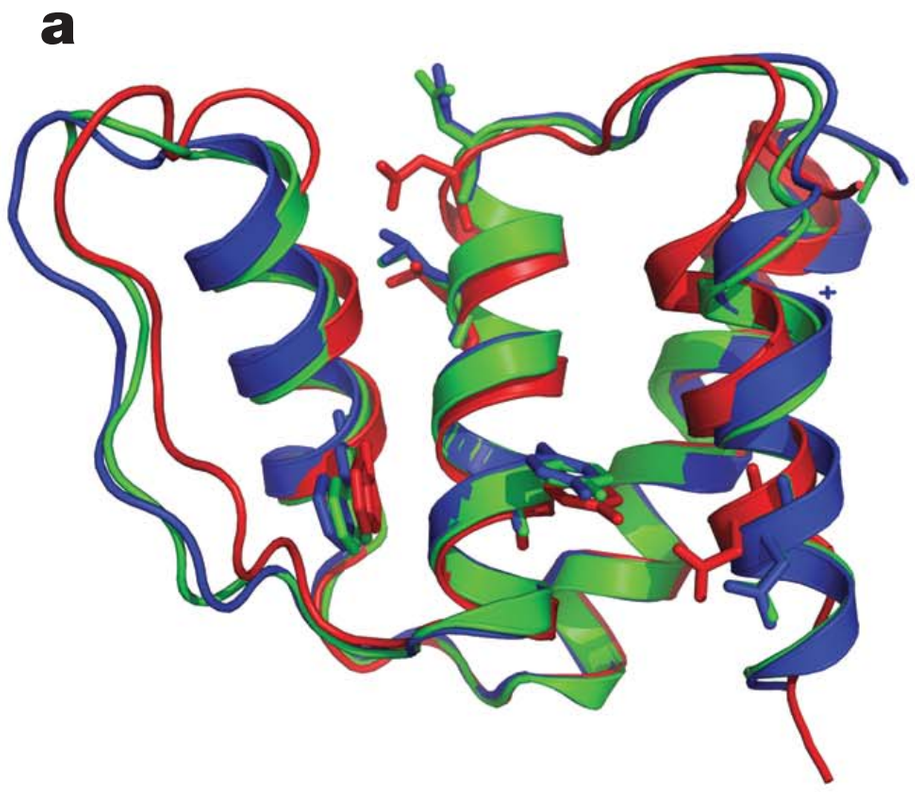
High-resolution structure prediction and the crystallographic phase problem

Bin Qian[1]*, Srivatsan Raman[1]*, Rhiju Das[1]*, Philip Bradley[1], Airlie J. McCoy[2], Randy J. Read[2] & David Baker[1]

The energy-based refinement of low-resolution protein structure models to atomic-level accuracy is a major challenge for computational structural biology. Here we describe a new approach to refining protein structure models that focuses sampling in regions most likely to contain errors while allowing the whole structure to relax in a physically realistic all-atom force field. In applications to models produced using nuclear magnetic resonance data and to comparative models based on distant structural homologues, the method can significantly improve the accuracy of the structures in terms of both the backbone conformations and the placement of core side chains. Furthermore, the resulting models satisfy a particularly stringent test: they provide significantly better solutions to the X-ray crystallographic phase problem in molecular replacement trials. Finally, we show that all-atom refinement can produce *de novo* protein structure predictions that reach the high accuracy required for molecular replacement without any experimental phase information and in the absence of templates suitable for molecular replacement from the Protein Data Bank. These results suggest that the combination of high-resolution structure prediction with state-of-the-art phasing tools may be unexpectedly powerful in phasing crystallographic data for which molecular replacement is hindered by the absence of sufficiently accurate previous models.

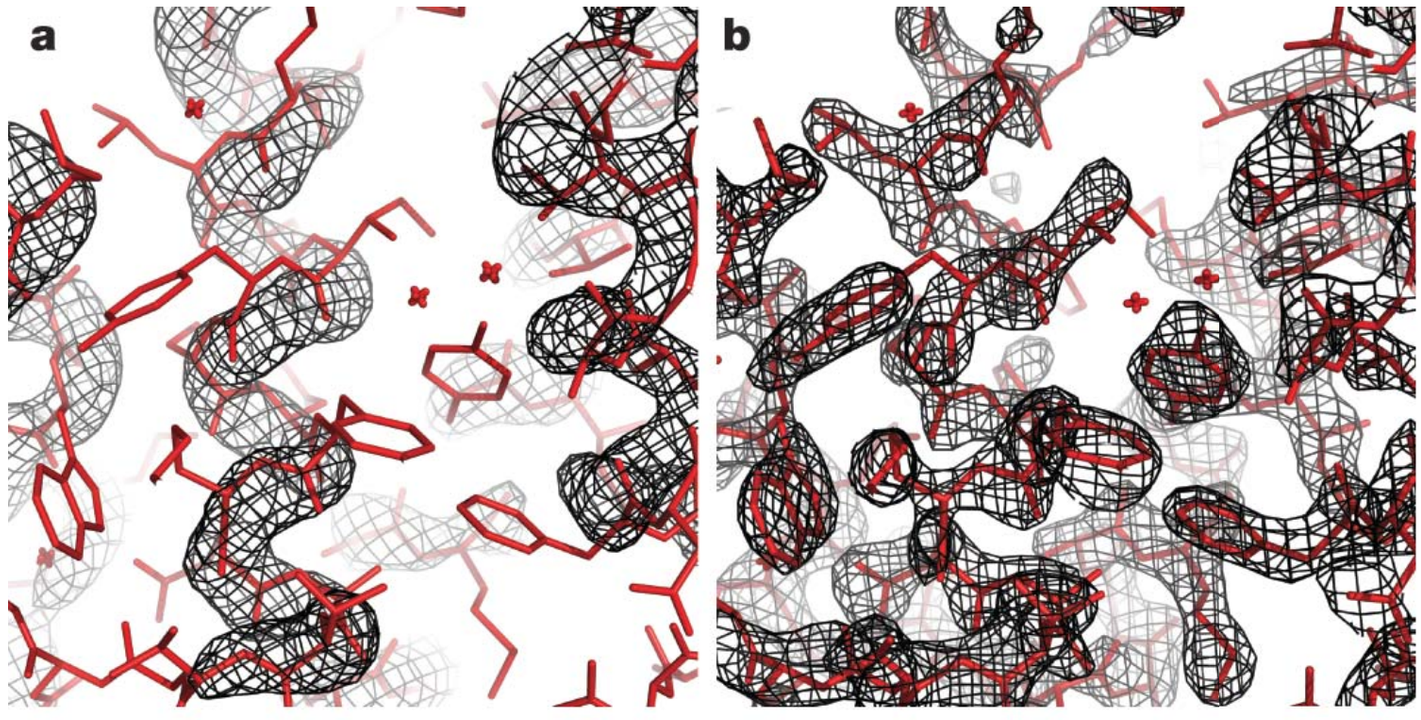


X-ray and NMR structure refinement contd.



Acyl CoA binding protein

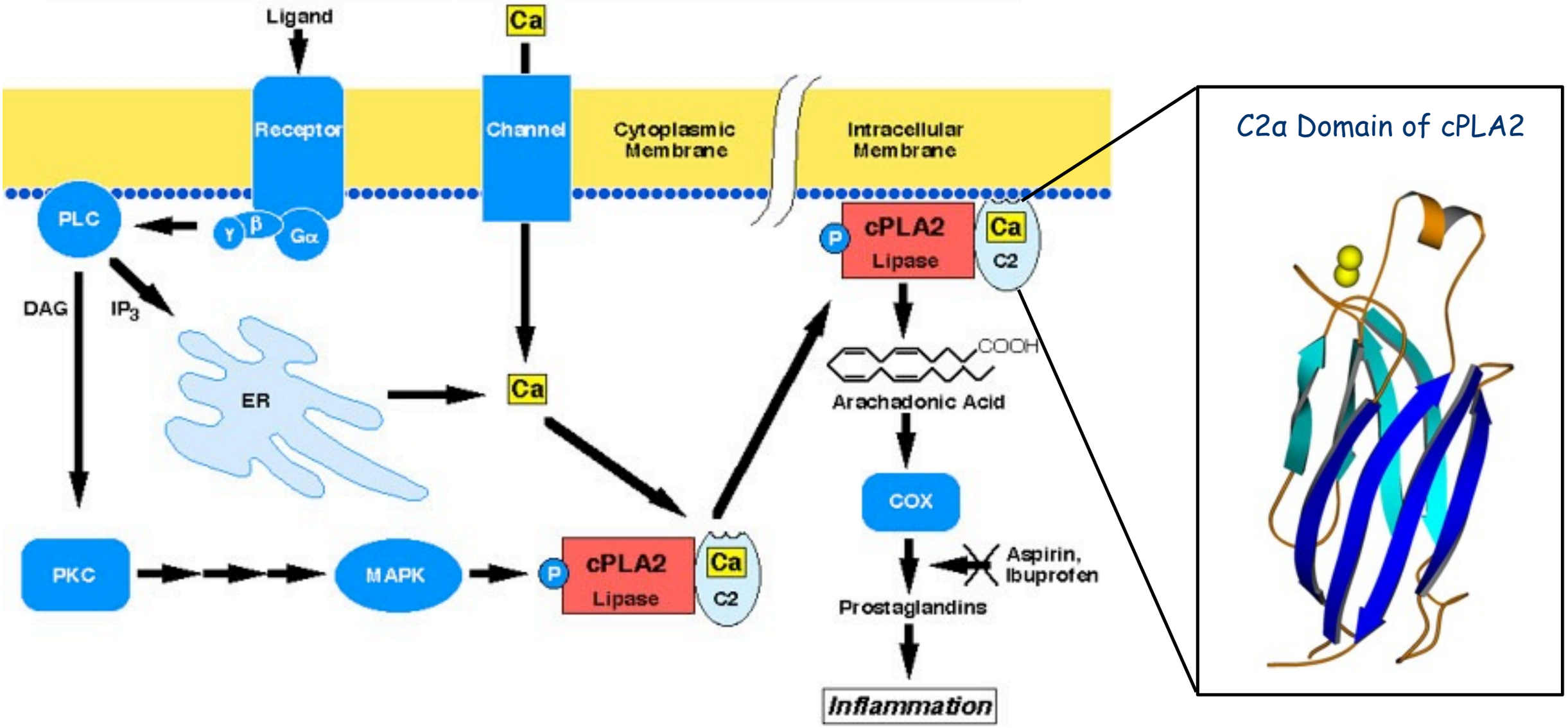
Superposed
crystal structure (blue),
NMR model (red) and
the lowest energy all-atom refined model (green)



Solving the X-ray crystallographic phase problem via molecular replacement: Improvement in electron density using model from rebuilding and refinement in molecular replacement searches.

Structure refinement

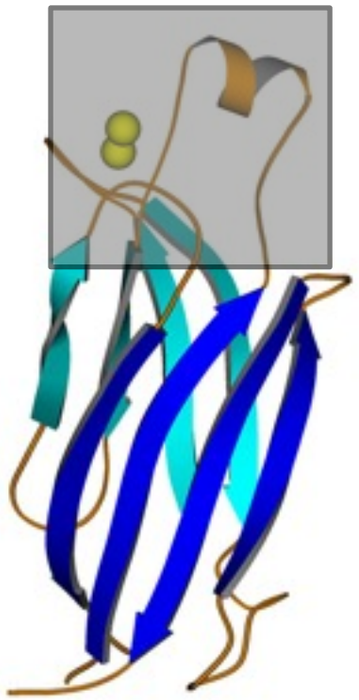
How does $cPLA_2$ propagate the Inflammation Signaling Pathway?



X-Ray Vs NMR

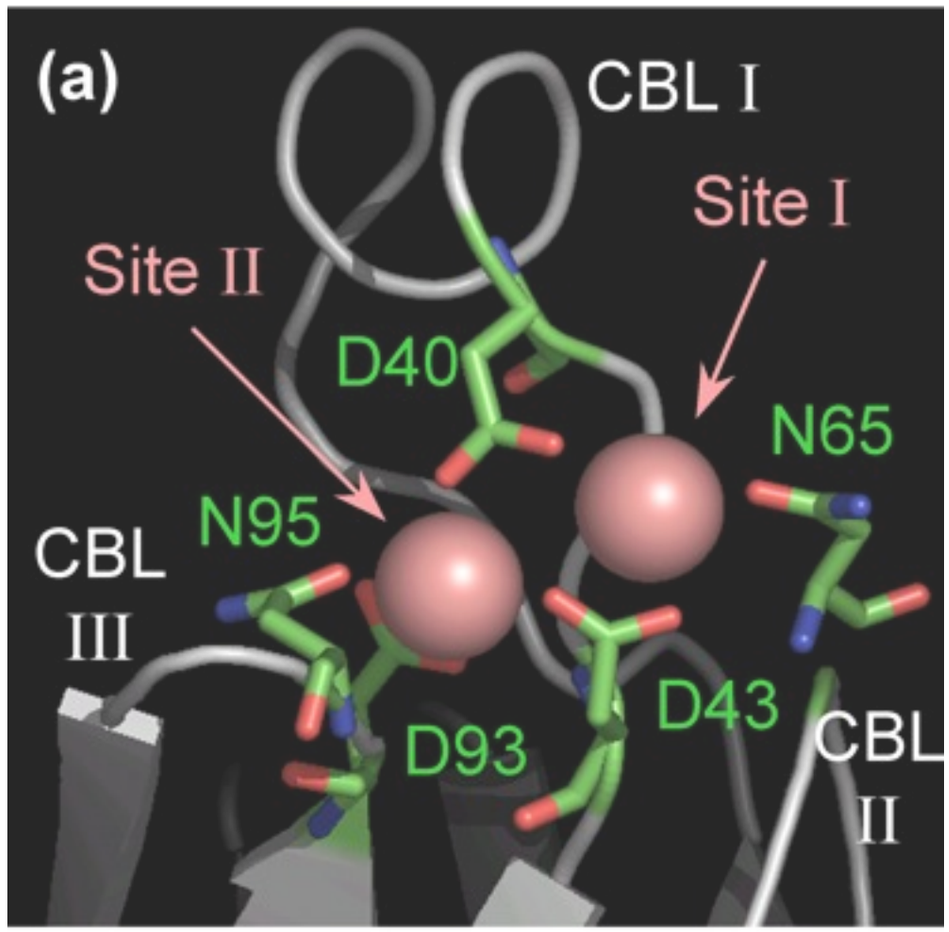
Used a **combined experimental** and **computational** approach to understand the differences in secondary structure

Membrane docking region



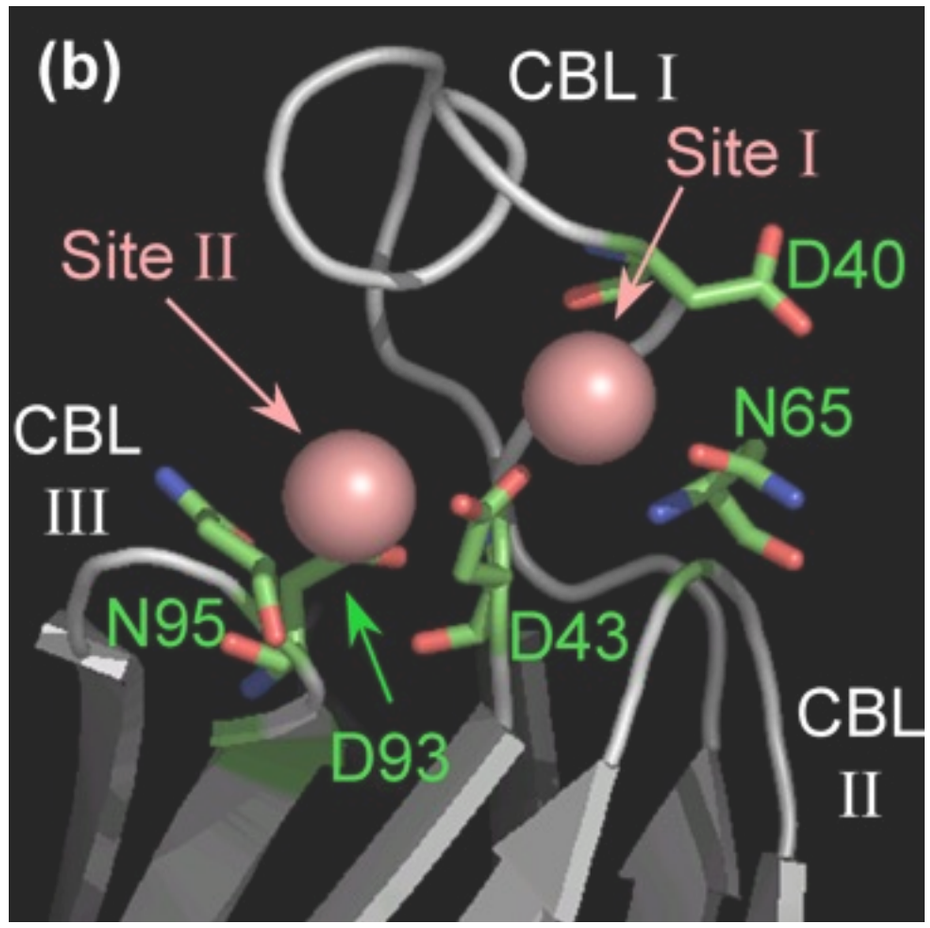
X-ray Structure

O. Perisic *et al.,* J Biol. Chem. 1998



NMR Structure

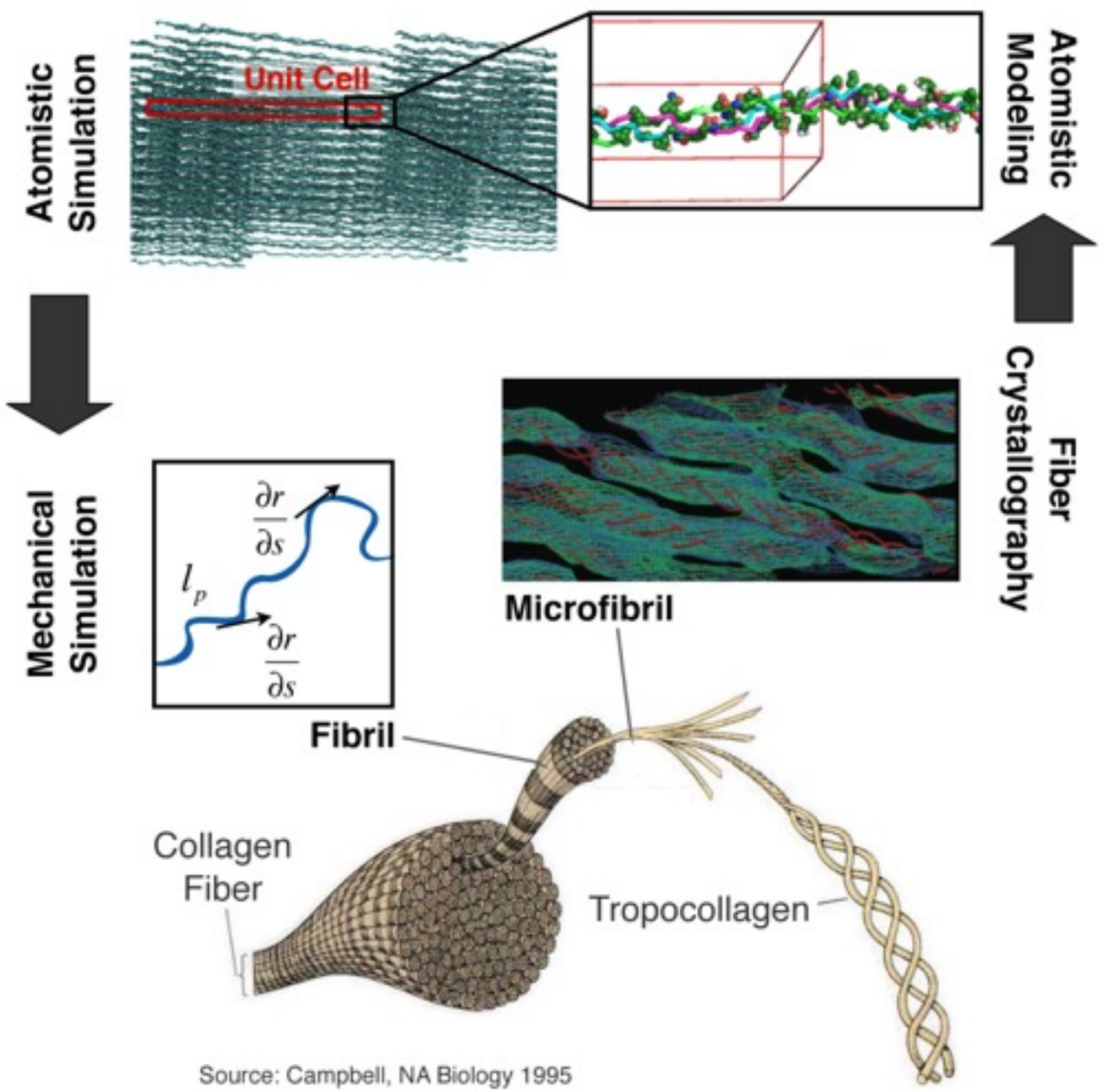
G. Xu *et al.*, J Mol. Biol. 1998



Malmberg, Varma, Jakobsson & Falke, Biochemistry. 2004
Varma & Jakobsson, Biophys. J. 2007

Post processing structural information

Collagen



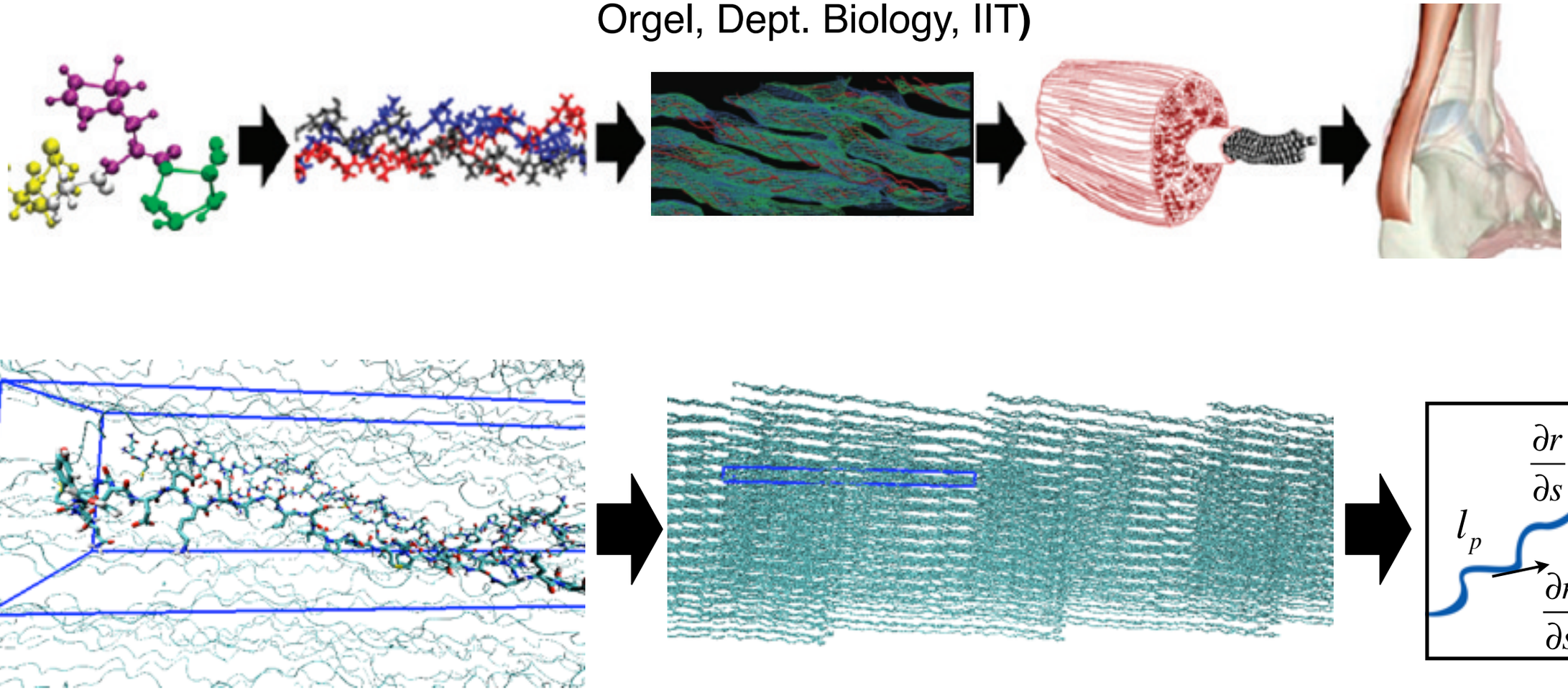
Source: Campbell, NA Biology 1995

Collagen cont.

Collagen is the main protein in our connective tissue

What gives collagen its strength?

Amino Acid → **Collagen Molecule** → **Collagen Microfibril** (X-ray structure, Orgel, Dept. Biology, IIT) → **Collagen Fiber** → **Tissue (tendon, bone)**

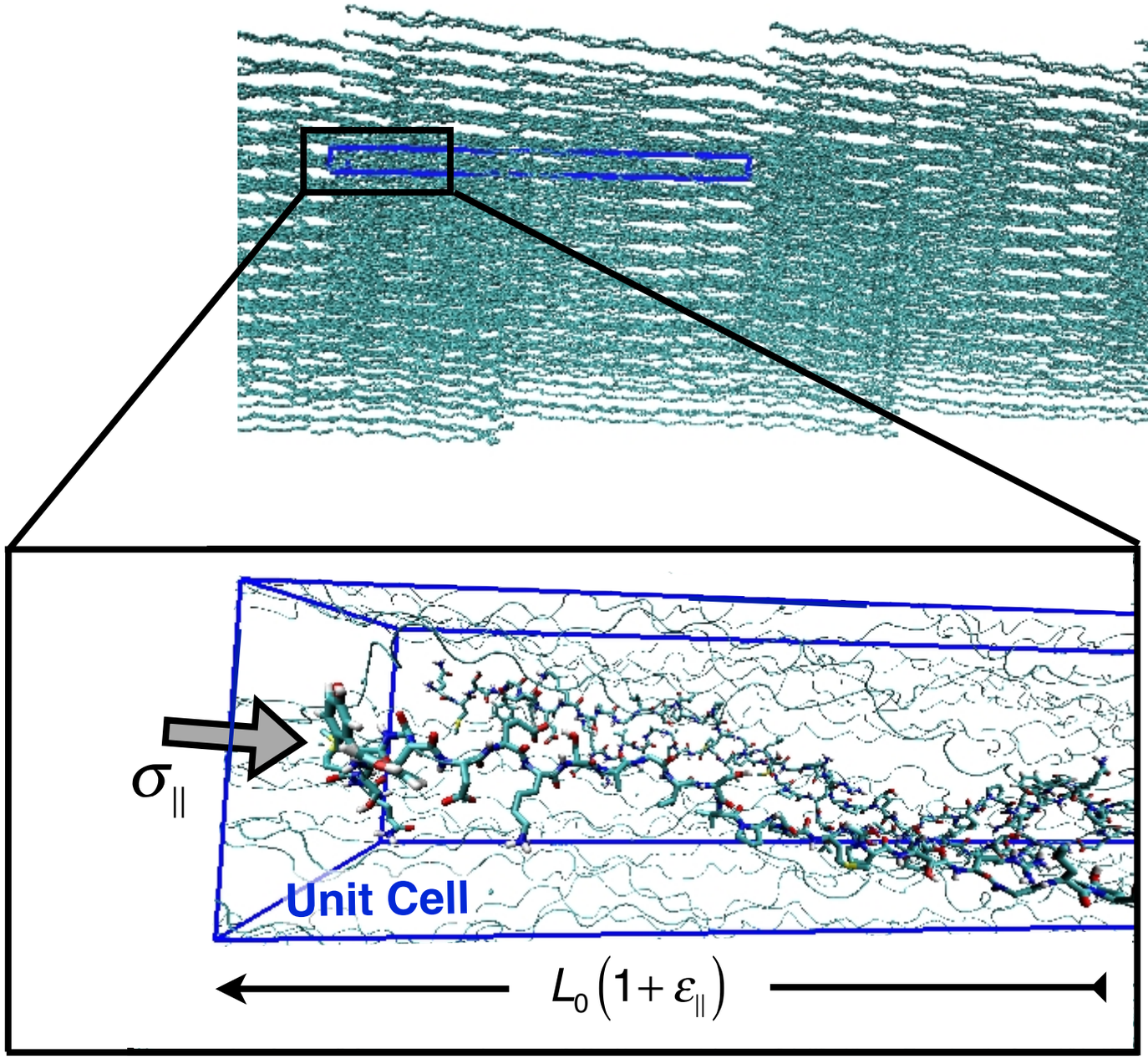


Atomistic Model → **Coarse-grained model** → **Mechanical model**

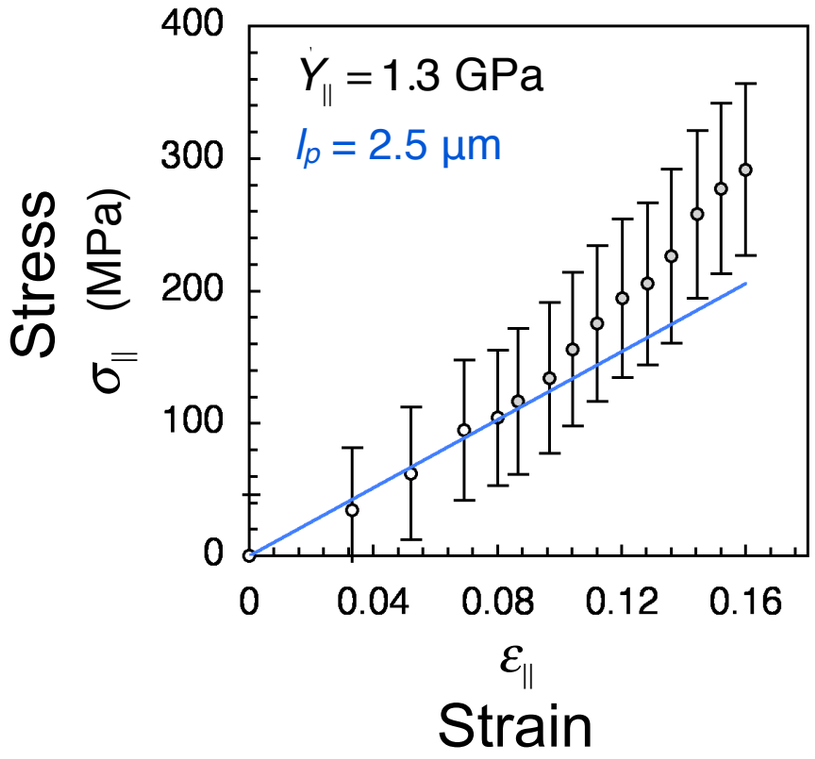
Cartoon: adapted from Gauitieri *et al.* 2010

Varma *et al.*, In Preparation

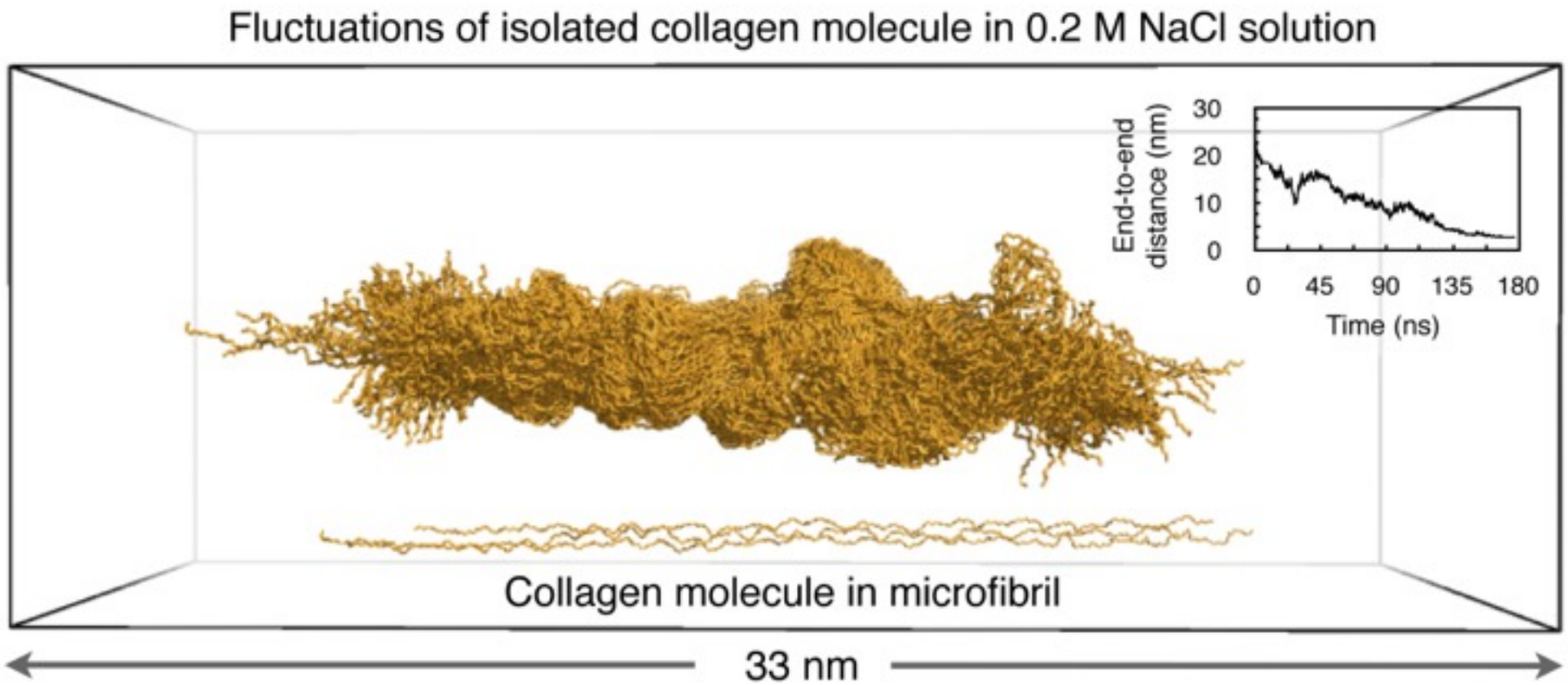
Collagen cont.

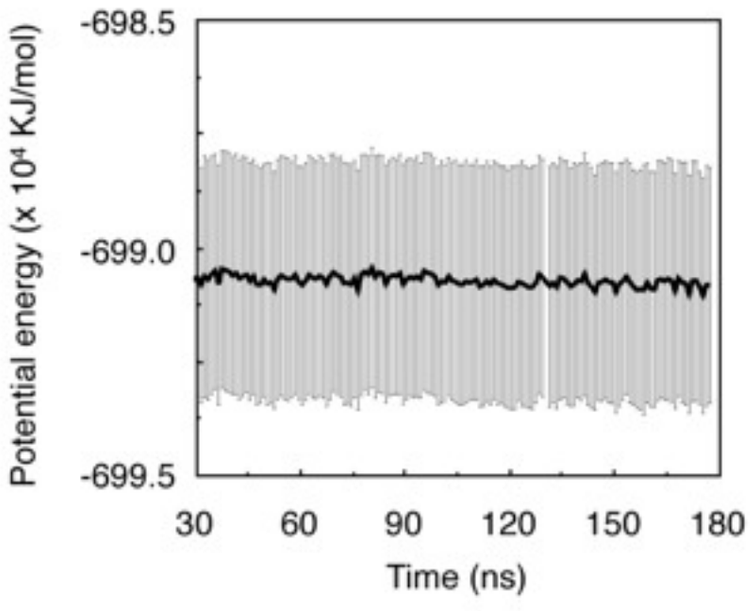


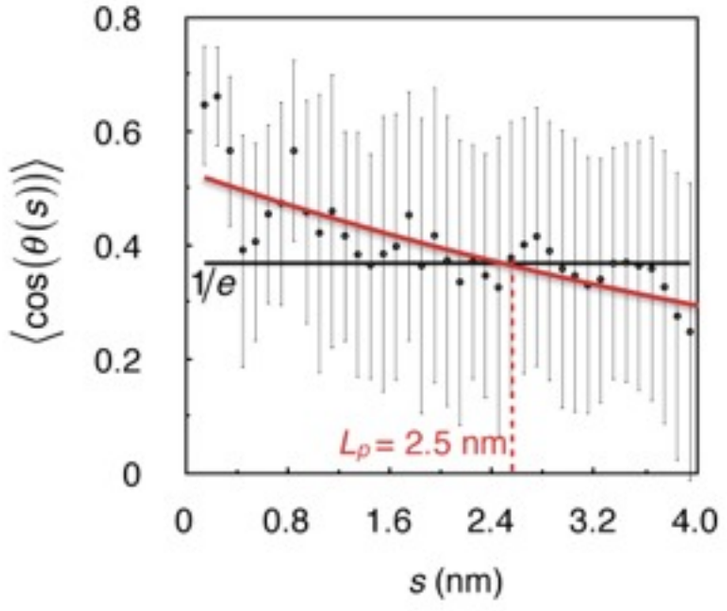
Y - Young's modulus
l_p - persistence length



Collagen cont.





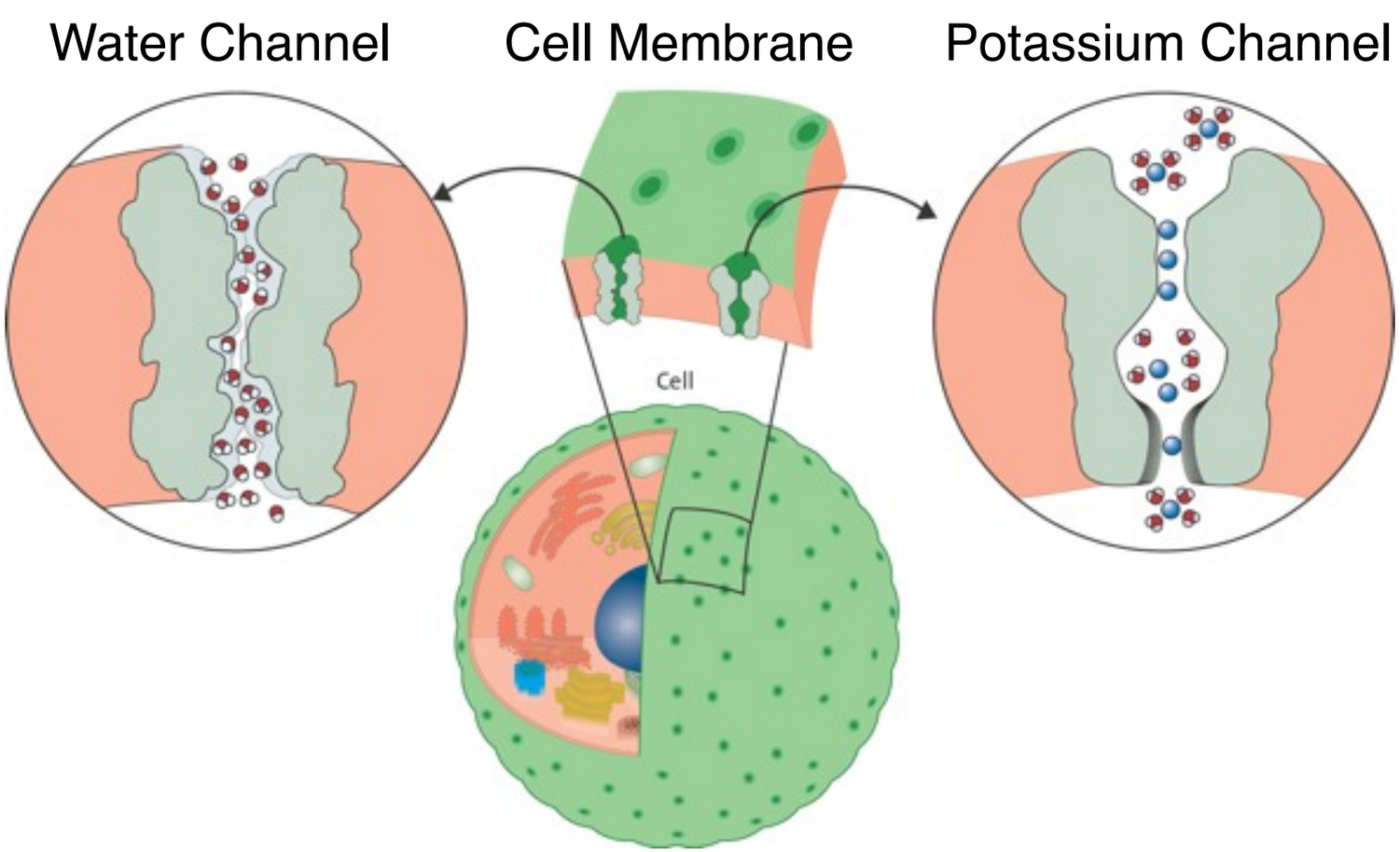


Potassium Channels

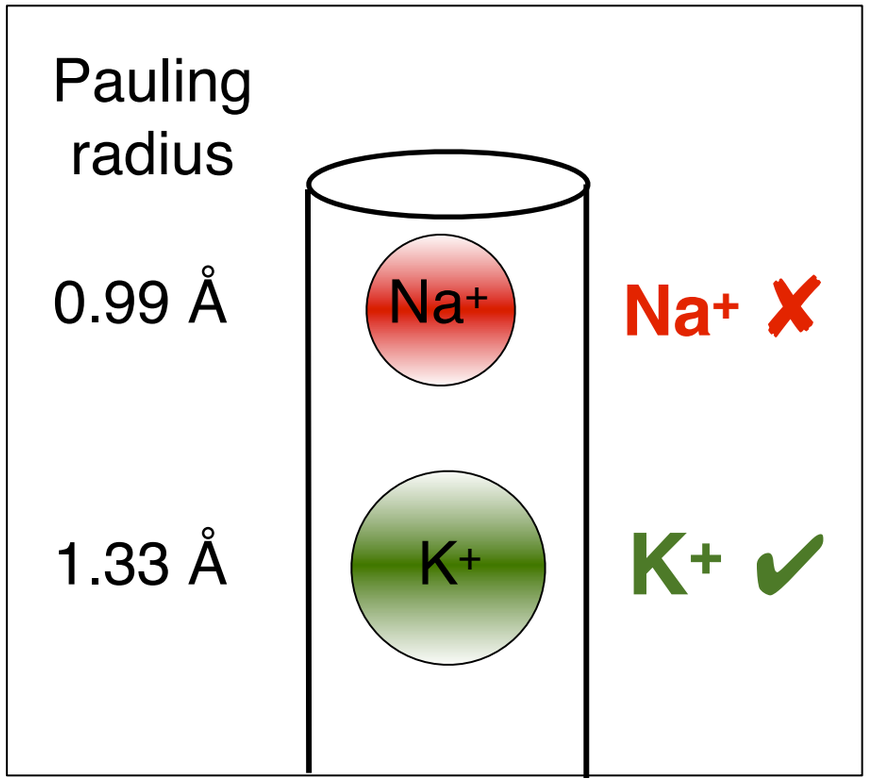
Selectively regulate K^+ concentration gradients across cell membranes to enable numerous physiological tasks, including nerve conduction and muscle contraction

What is the basis for K^+/Na^+ selectivity?

Mackinnon - Nobel Prize in 2003

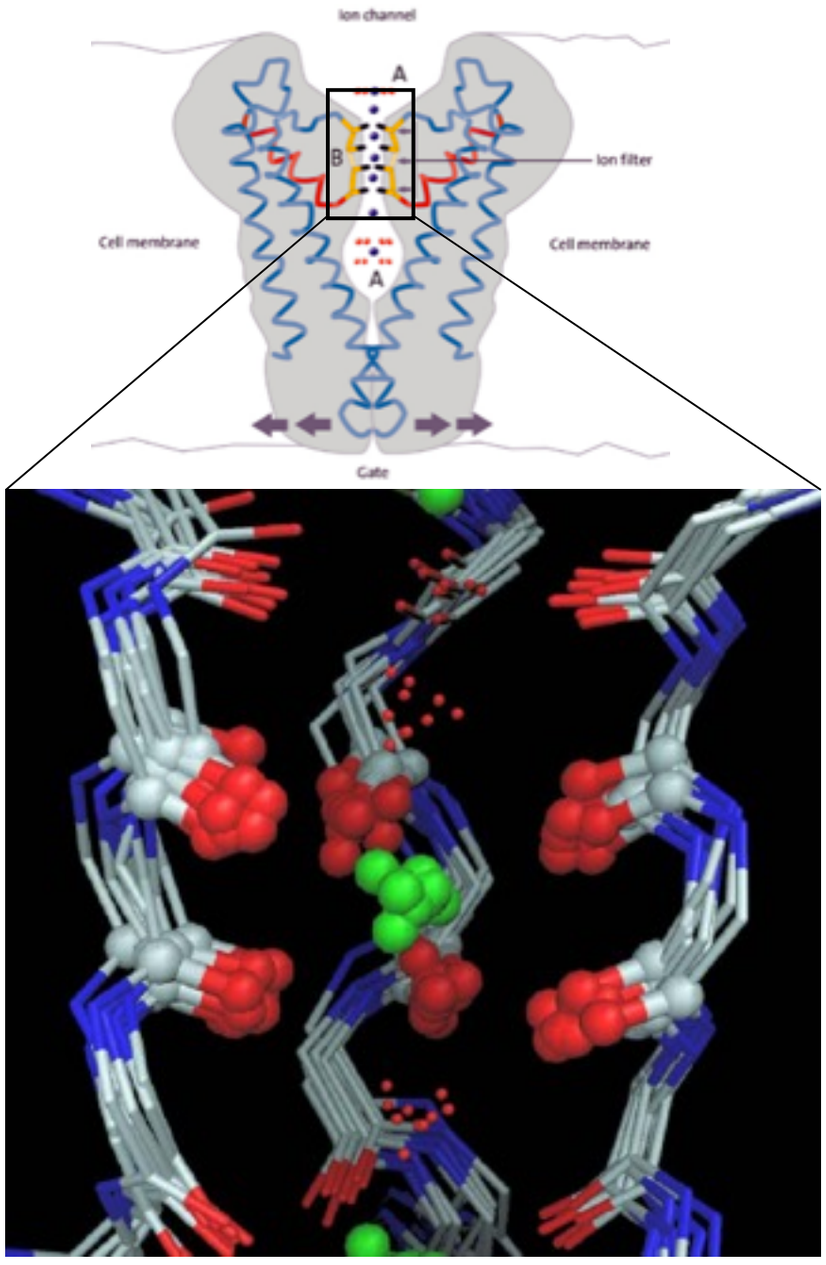


Cartoon taken from: http://nobel.se



Problems with the conventional picture

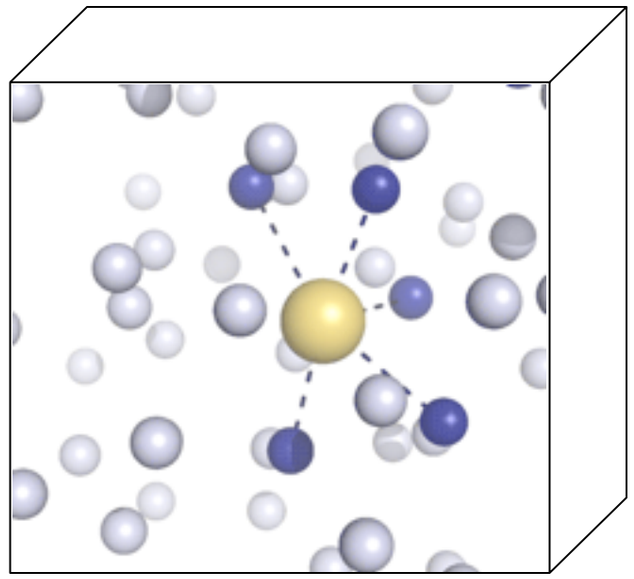
Thermal fluctuations obscure sub-angstrom size differences between Na^+ and K^+ ions



Noskov *et al.,* Nature 2004

New picture of K^+ in water

Quantum mechanical simulation



The probability to find an 8-fold coordination is negligible in liquid water

The binding sites in K-channels

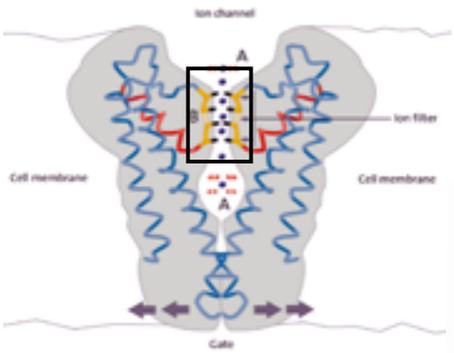
(1) DO NOT mimic the structure of K^+ ions in bulk water and, in fact,

(2) over-coordinate the K^+ ion.

Varma & Rempe, Biophys. Chem., 2006

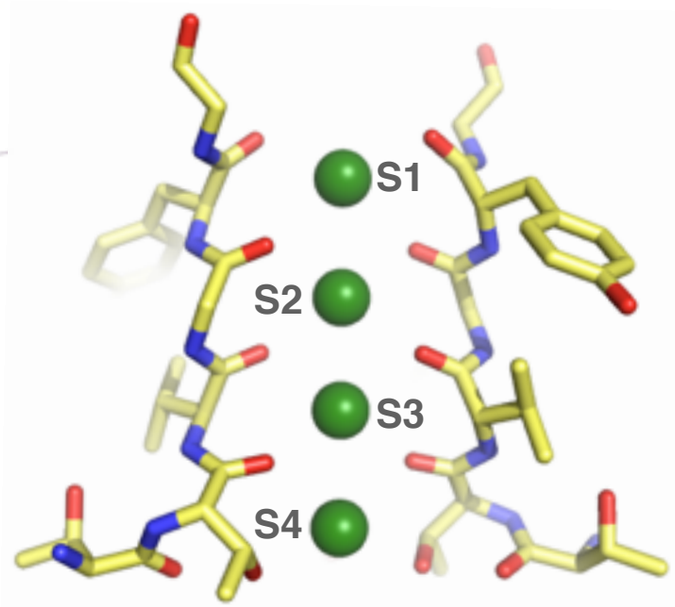
Problems with the conventional picture

K-channel filter adopts multiple configurations during ion binding

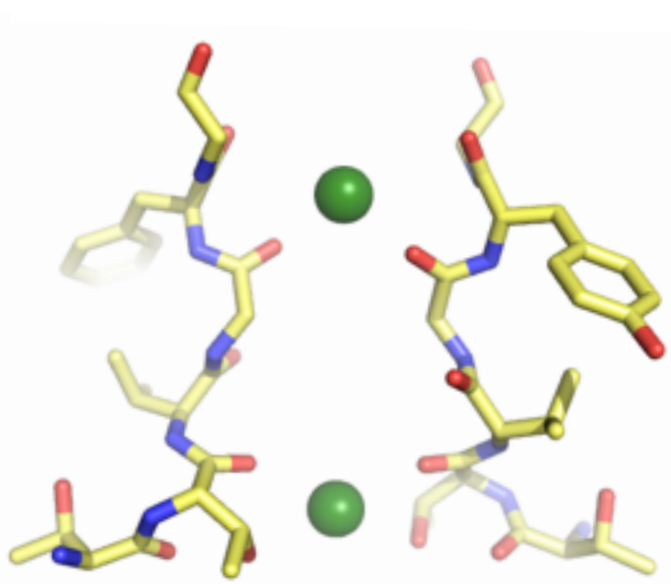


X-ray structures:
Mackinnon and coworkers

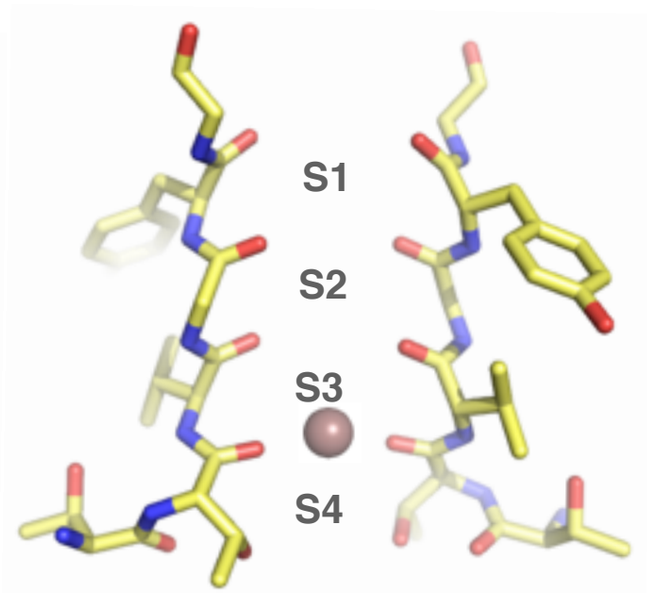
K^+ bound in
High K^+ concentration



K^+ bound in
Low K^+ concentration

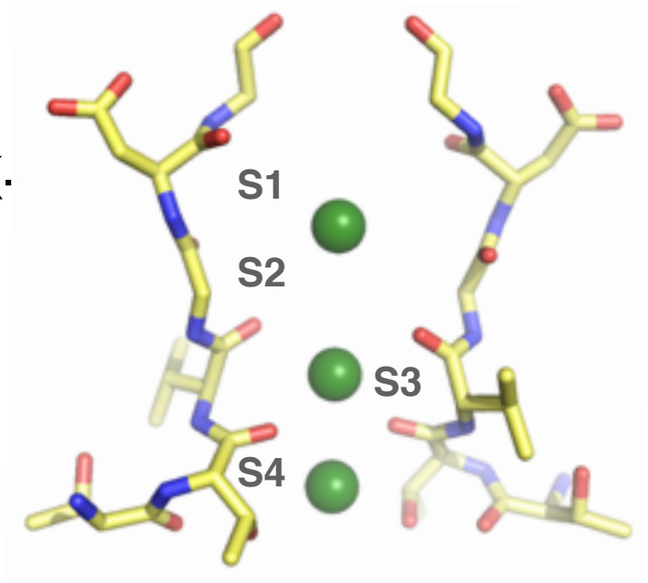


Na^+ bound in
Low K^+ concentration

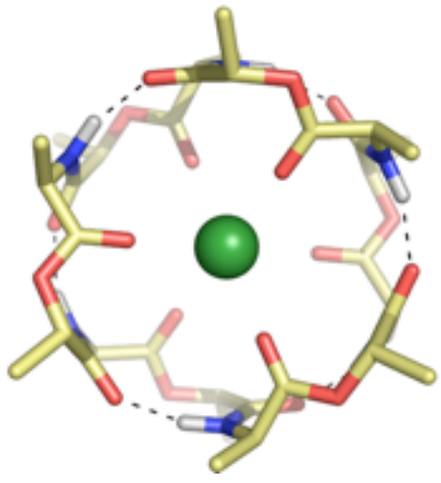


NaK channels
share an overall architecture with K-channels but are non-selective

X-ray structure:
Jiang and coworkers



Valinomycin
does not share an overall architecture with K-channels, but is as selective as K-channels



Different mechanisms of K^+/Na^+ selectivity

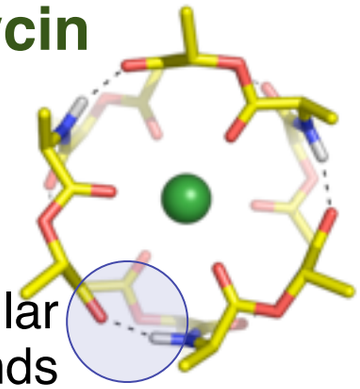
Na^+ and K^+ in liquid water (Low coordination)

Introduce special ligand chemistry

Liquid Dimethyl Thioformamide

Constrain cavity size

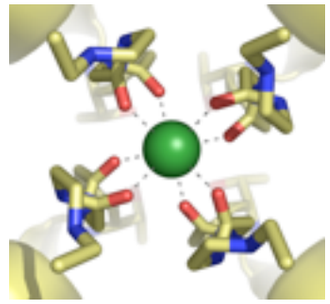
Valinomycin



Intra-molecular H-bonds

Constrain coordination number > 6

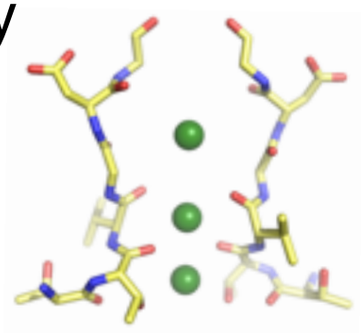
Strongly selective K-channels



Weaken constraint on high coordination

Enhance Flexibility

Weakly selective NaK channel



Additional possibilities

Different mechanisms of K^+/Na^+ selectivity

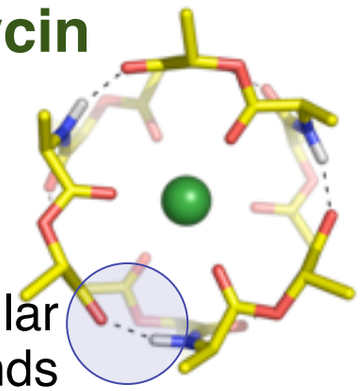
Na^+ and K^+ in liquid water (Low coordination)

Introduce special ligand chemistry

Liquid Dimethyl Thioformamide

Constrain cavity size

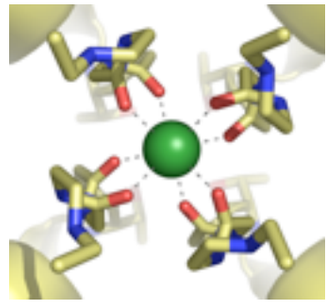
Valinomycin



Intra-molecular H-bonds

Constrain coordination number > 6

Strongly selective K-channels



Additional possibilities

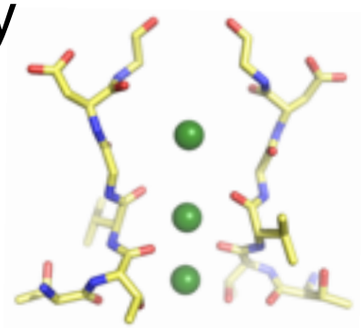
Weaken constraint on high coordination

Introduce H-bond donors in local env.

Weakly selective K-channels (Eg. HERGs)

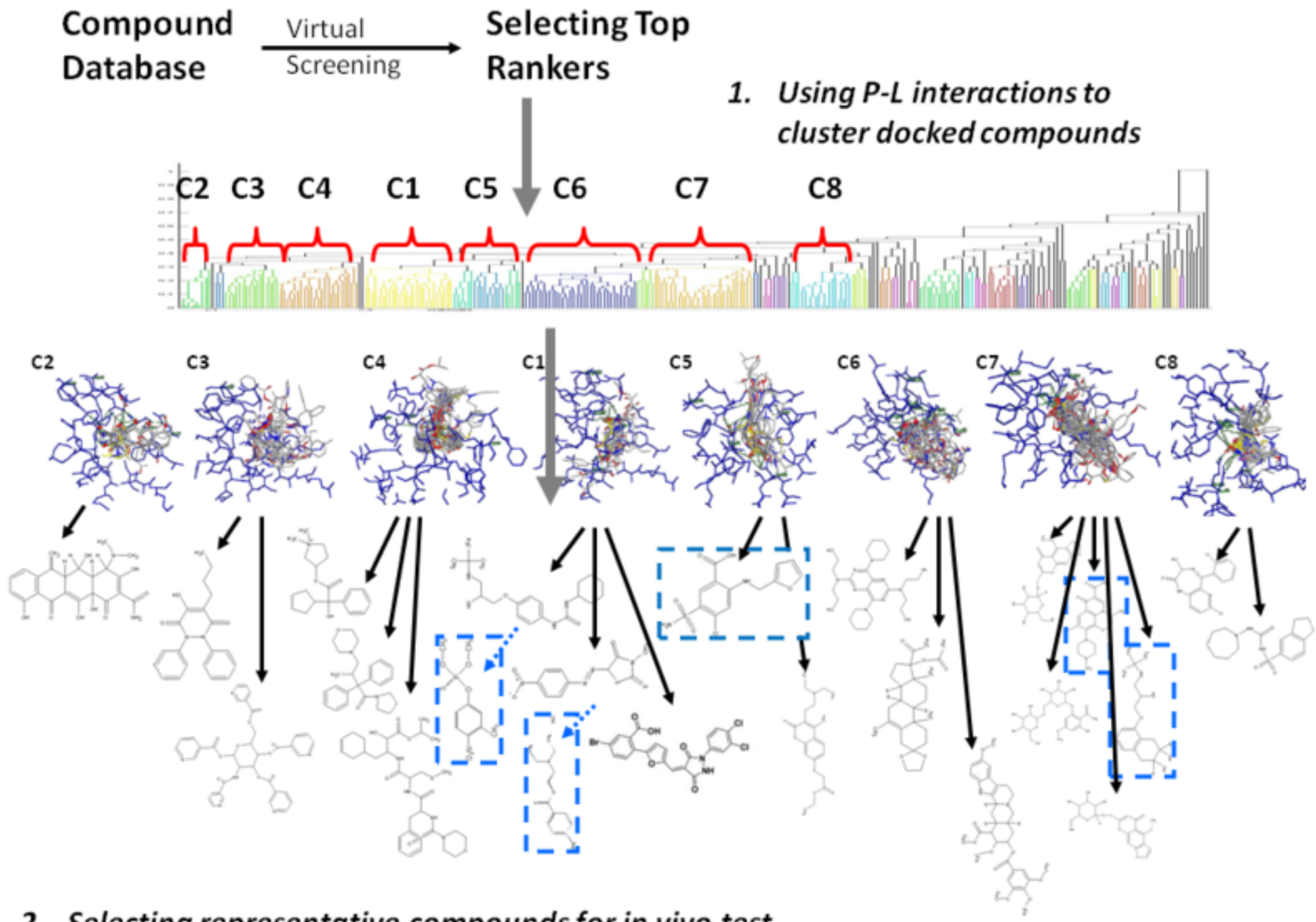
Enhance Flexibility

Weakly selective NaK channel



Varma *et al*.
JBC 2006
BJ 2007
JMB 2008
JACS 2008
BJ 2010
JGP 2011

Computer-aided drug design



2. *Selecting representative compounds for in vivo test*