Foundations of biomaterials: Models of protein solvation

L. Ridgway Scott

The Institute for Biophysical Dynamics,
The Computation Institute, and the
Departments of Computer Science and Mathematics,
The University of Chicago

Solvation in picomaterials

Polar liquids play singular role in

- Protein interactions (water most important molecule in biology)
- Ionic solvation (ionic liquids)
- Environmental pollution (HCl hydrate)
- Energy production (methane hydrate)
- Batteries
- Photovoltaic cells

Dielectric effect critical for protein interactions

Picoscale models needed

Modeling issues for solvation

- Solvents are mobile, not fixed in orientation
- Nonlocal effects (frequency dependence): high dimension
- Nonlinear models: do they achieve same results as nonlocal ones?
- Ionic effects: can size effects be modeled via continuum equations? E.g., Na Cl
- Algorithms to resolve nonlinear (e.g., ionic) interactions
- Need molecular scale models of solvation

Electrostatic modeling: ionic effects

Protein sidechains have large electrostatic gradients

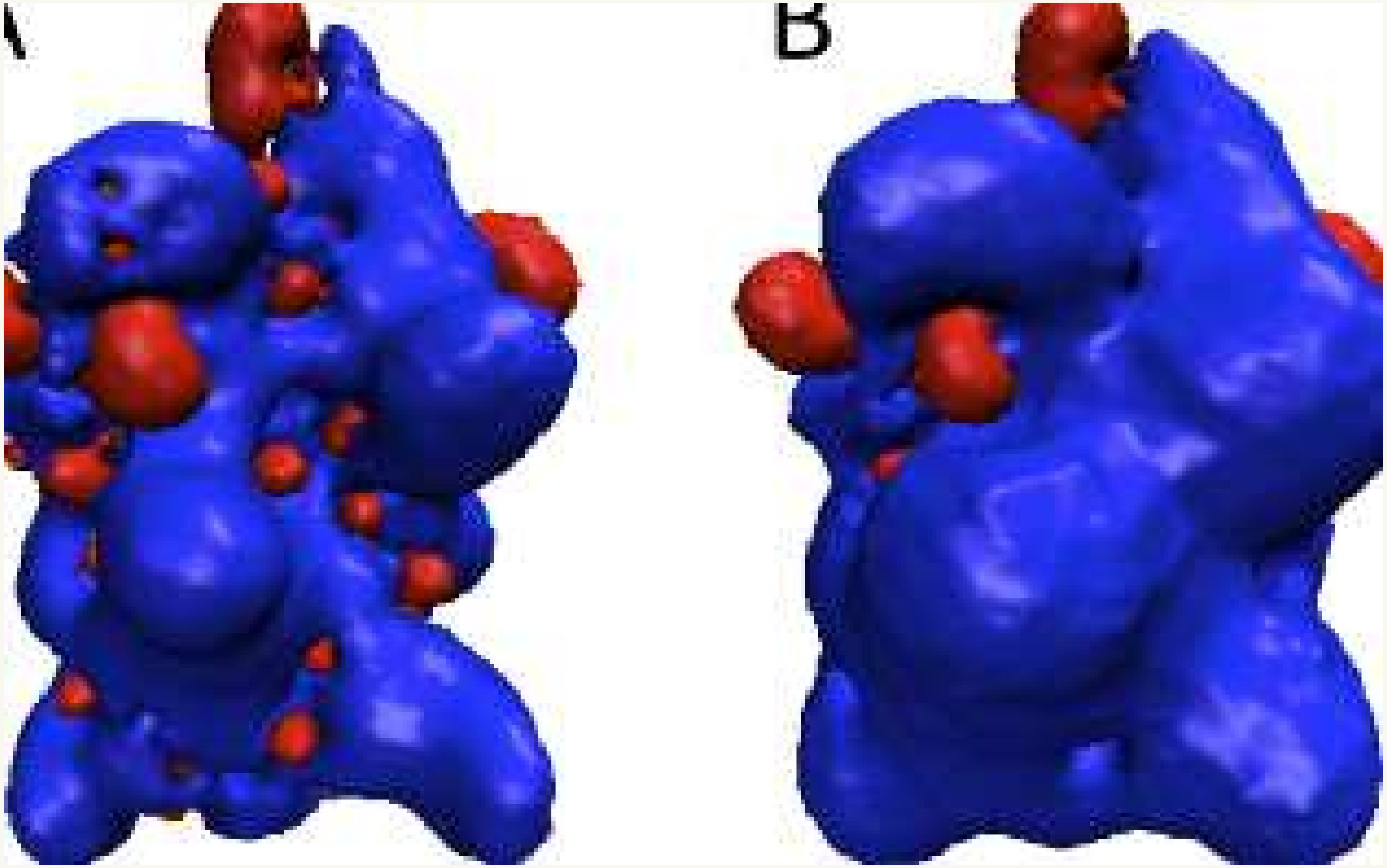
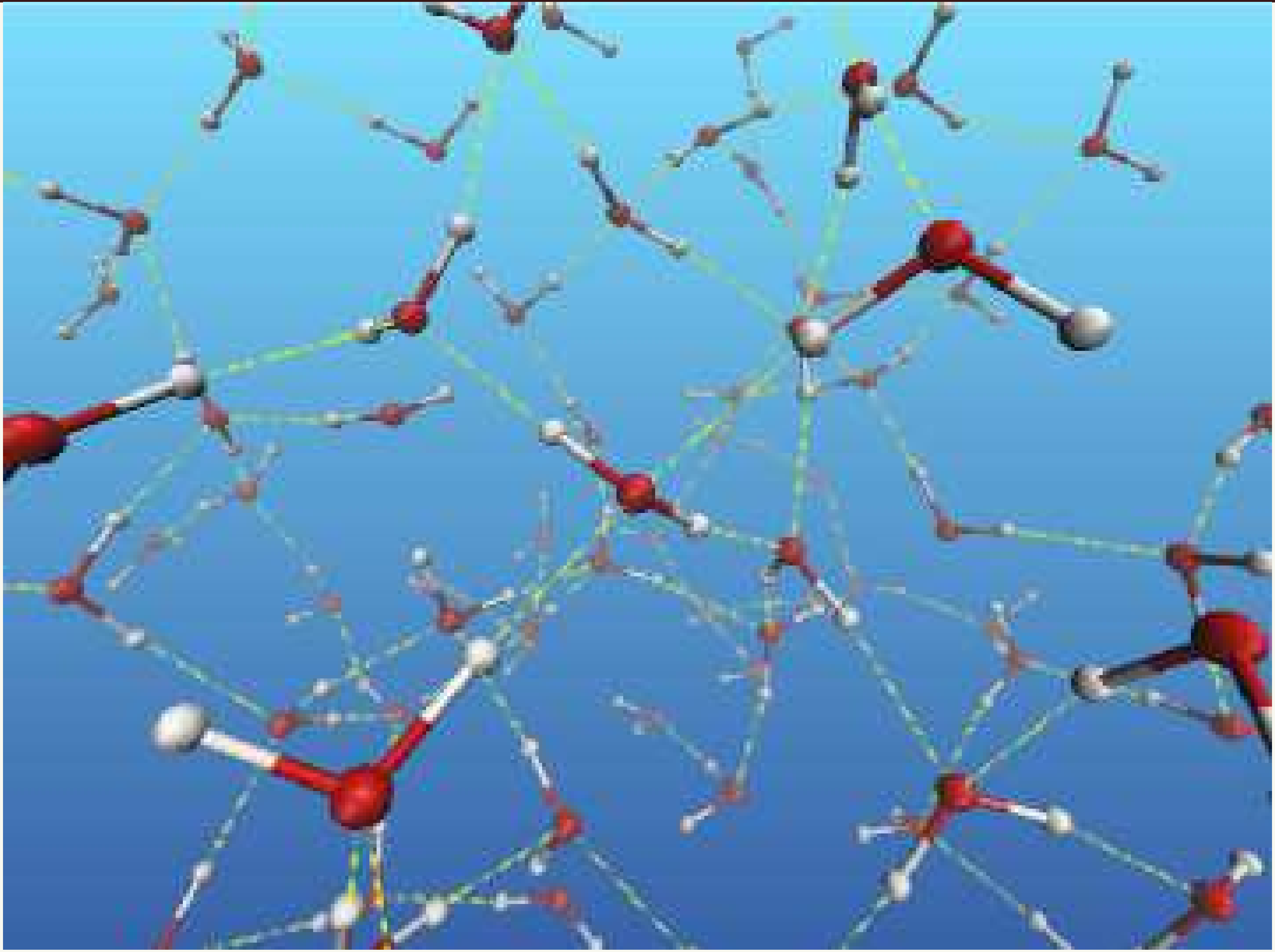


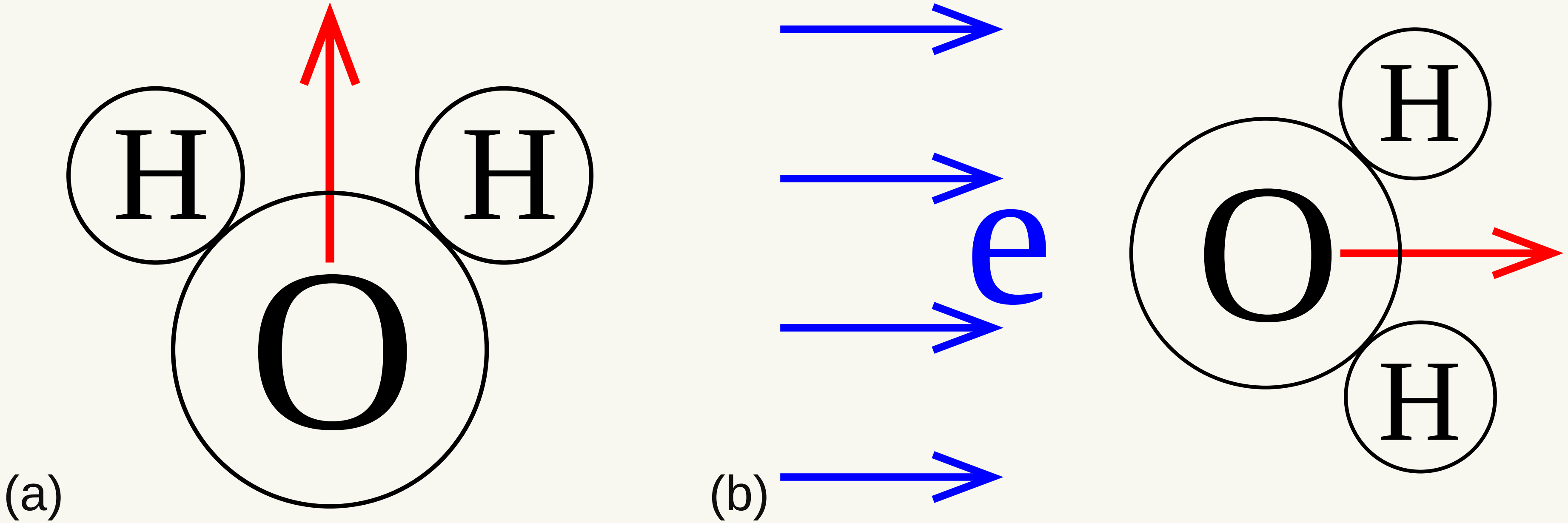
Figure 1: Different models suggest different modes of interactions. Shown are two nonlocal dielectric models with (A) and without (B) ionic effects.

Water is complicated



Water network from a molecular dynamics simulation

Dielectric effect of water



(a) Polarity of water molecule

(b) Ability to rotate allows water to align with electric field e

Result: water screens electric field

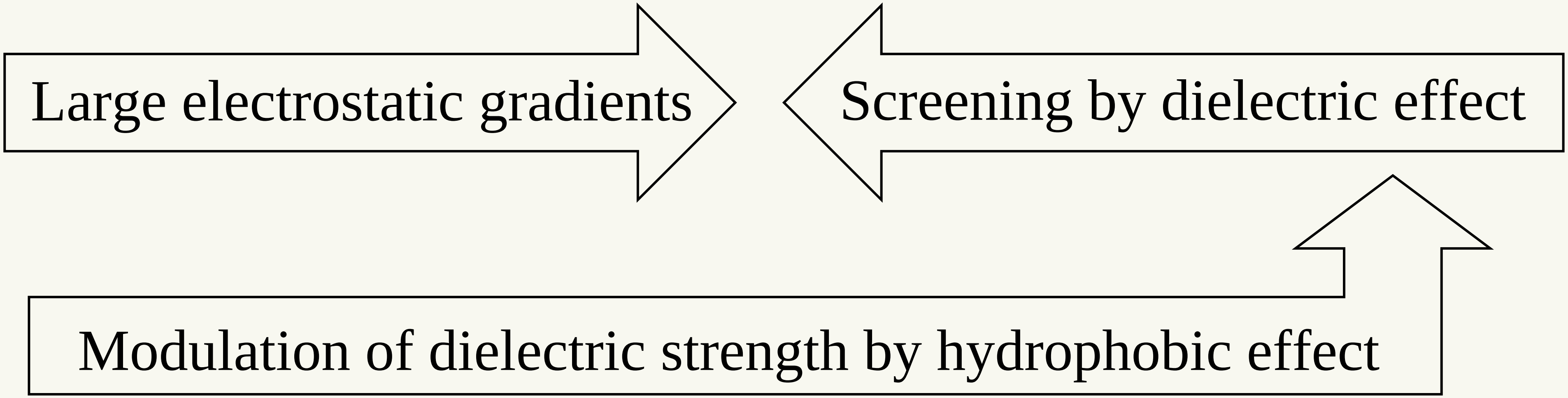
Competing effects

Competing effects: why this is so hard

Protein sidechains have large electrostatic gradients

Water is a strong dielectric

Hydrophobic groups modify the water structure



Large electrostatic gradients

Screening by dielectric effect

Modulation of dielectric strength by hydrophobic effect

Figure 2: Three competing effects that determine protein behavior. These conspire to weaken interactive forces, making biological relationships more tenuous and amenable to mutation.

Charges in a dielectric

Charges in a dielectric are like lights in a fog.

Dielectric model

Consider two charge distributions ρ (fixed charges) and γ (polar groups free to rotate). Resulting electric potential ϕ satisfies

$$\Delta\phi = \rho + \gamma, \tag{1}$$

where the dielectric constant of free space is set to one.

Write $\phi = \phi_\rho + \phi_\gamma$, where $\Delta\phi_\gamma = \gamma$ and $\Delta\phi_\rho = \rho$.

Ansatz of Debye [3]: the electric field $\mathbf{e}_\gamma = \nabla\phi_\gamma$ is parallel to (opposing) the resulting electric field $\mathbf{e} = \nabla\phi$:

$$\nabla\phi_\gamma = (1-\varepsilon)\nabla\phi. \tag{2}$$

Thus $\nabla\phi_\rho = \nabla\phi - \nabla\phi_\gamma = \varepsilon\nabla\phi$ and

$$\nabla \cdot (\varepsilon\nabla\phi) = \rho. \tag{3}$$

Polarization field and Debeye's Ansatz as projection

Define $\mathbf{p} = \nabla\phi_\gamma$: called the polarization field.

Recall $\mathbf{e} = \nabla\phi$.

Write $\mathbf{p} = (\epsilon - \epsilon_0)\mathbf{e} + \zeta\mathbf{e}^\perp$, so that

$$\epsilon = \epsilon_0 + \frac{\mathbf{p}\cdot\mathbf{e}}{\mathbf{e}\cdot\mathbf{e}},$$

with the appropriate optimism that $\mathbf{p} = 0$ when $\mathbf{e} = 0$.

That is, $\epsilon - \epsilon_0$ reflects the correlation between $\mathbf{p}$ and $\mathbf{e}$.

As defined, ϵ is a function of $\mathbf{r}$ and t, and potentially singular.

However, Debye postulated that a suitable average $\widetilde{\epsilon}$ should be well behaved:

$$\widetilde{\epsilon} = \epsilon_0 + \left\langle \frac{\mathbf{p}\cdot\mathbf{e}}{\mathbf{e}\cdot\mathbf{e}} \right\rangle.$$

Interpretation of ε

In bulk water ε is a (temperature-dependent) constant:

$$\varepsilon \approx 87.74 - 40.00\,\tau + 9.398\,\tau^2 - 1.410\,\tau^3, \quad \tau \in [0,1], \tag{4}$$

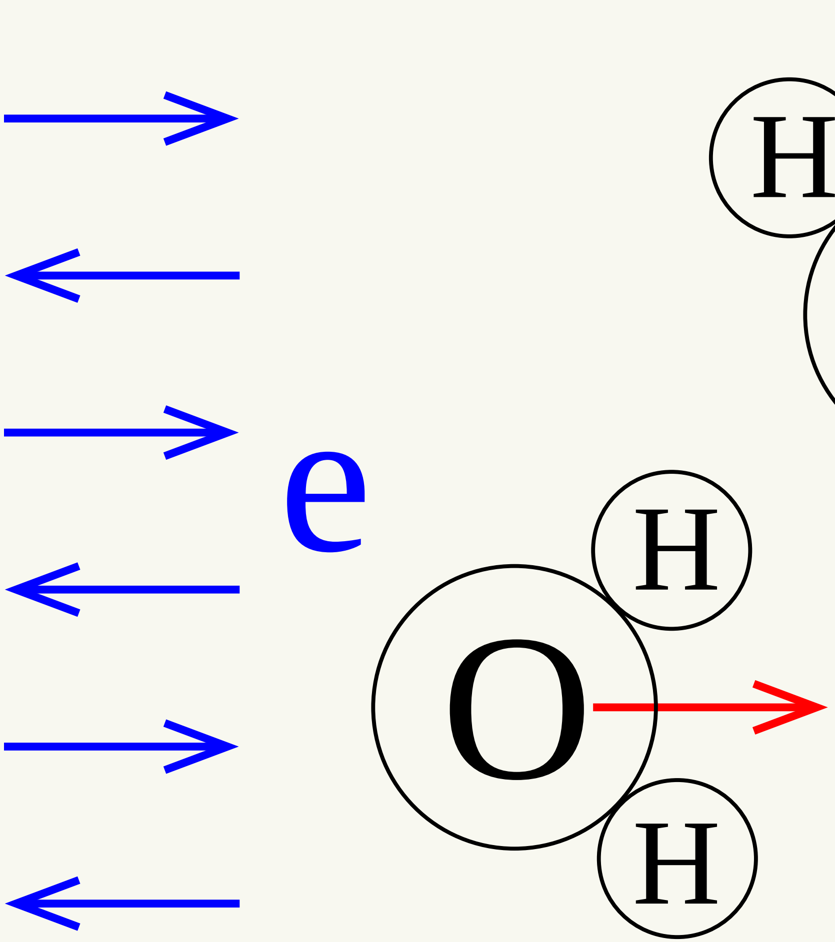
where $\tau = T/100$ and T is temperature in Centigrade (for $T > 0$) [4].

$\varepsilon >> 1$: opposing field strength $E_\gamma = \nabla\phi_\gamma$ much greater than inducing field.

ε increases with decreasing temperature;
when water freezes, it increases further:
for ice at zero degrees Centigrade, $\varepsilon \approx 92$.

Increased coherence yields increased dielectric

Model failure



But model fails when the spatial frequencies of the electric field $e = \nabla\phi$ are commensurate with the size of a water molecule, since the water molecules cannot orient appropriately to align with the field.

Model fixes

Manipulations leading to (3) valid when ε is an operator, even nonlinear.

Frequency-dependent versions of ε have been proposed, and these are often called ‘nonlocal’ models.

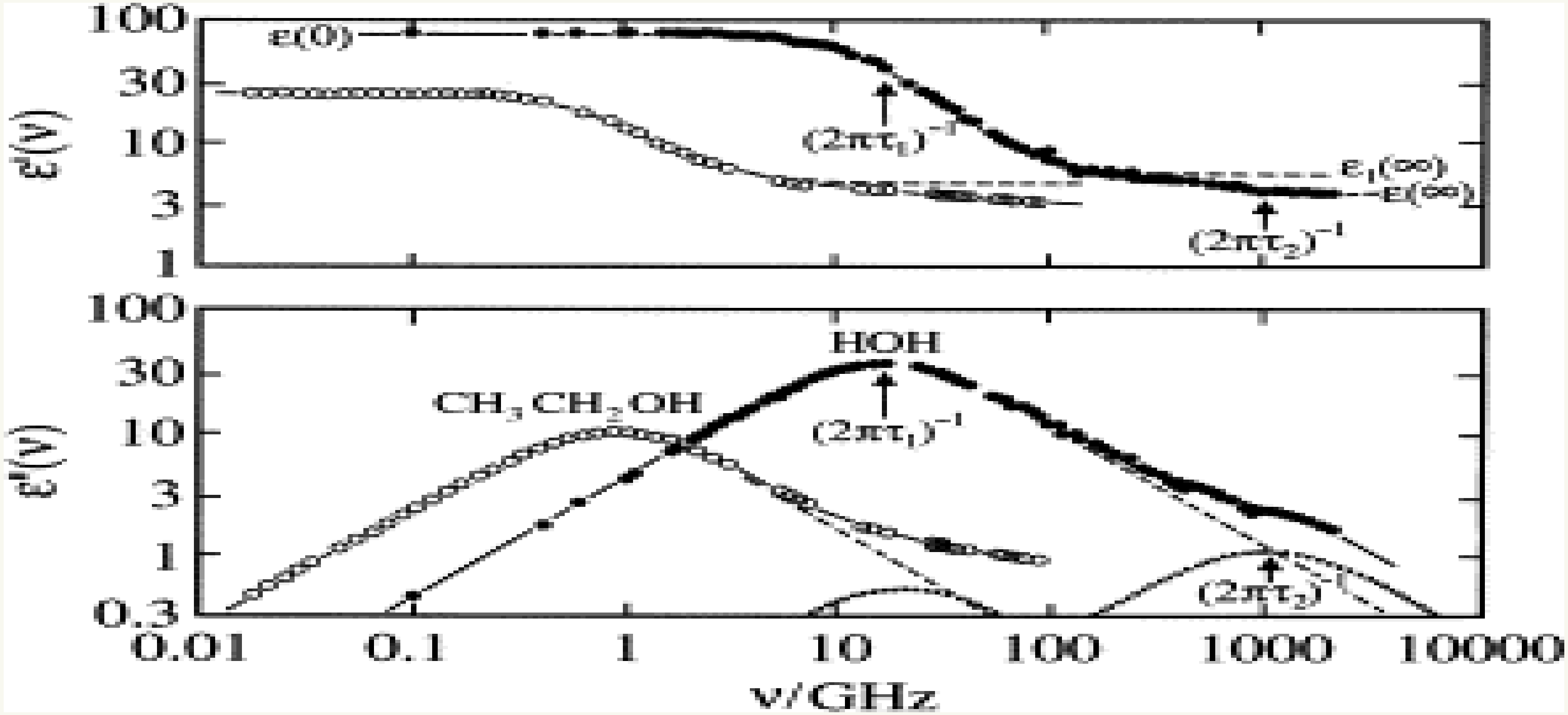
The operator ε must be represented either as a Fourier integral (in frequency space), or as an integral in physical space with a nonlocal kernel [1, 6].

Frequency dependence of dielectric constant

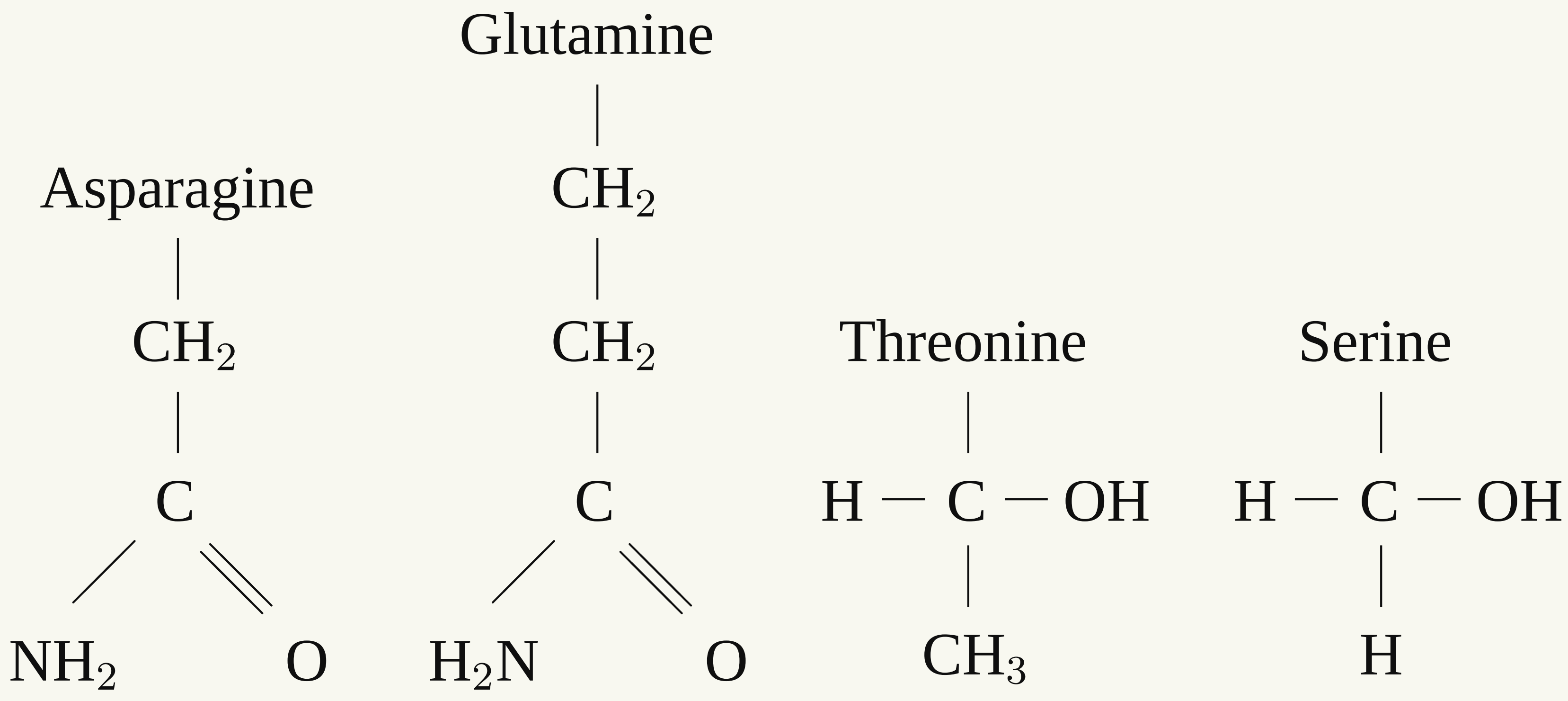
Debye observed that the effective permittivity is frequency dependent:

$$\epsilon(\nu) = \epsilon_0 + \frac{\epsilon_1 - \epsilon_0}{1 + \tau_D^2 \nu^2} \tag{5}$$

where τ_D is characteristic time associated with dielectric material and ν is temporal wave number. Many experiments have verified this [5]:

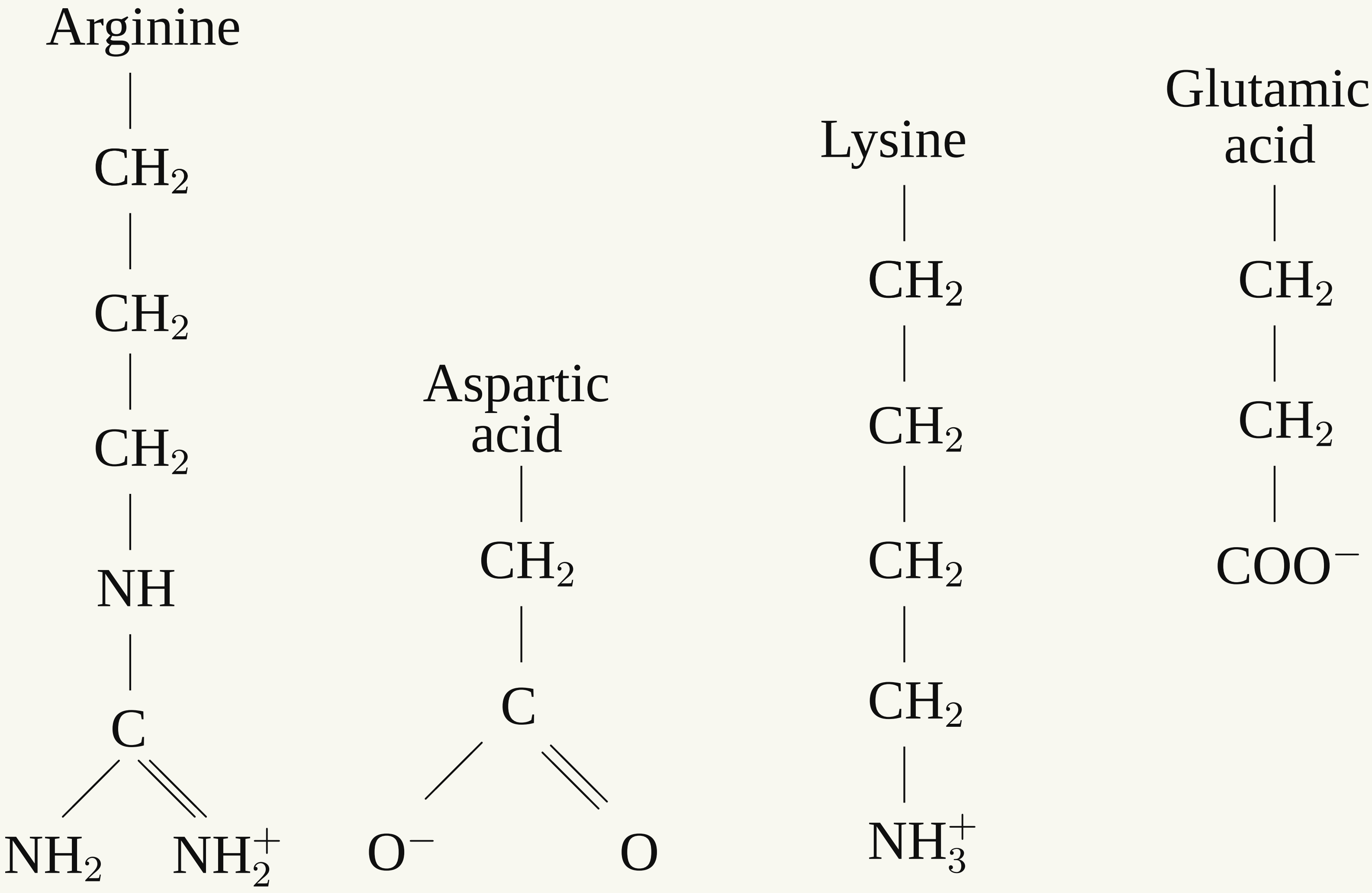


Polar residues cause spatial high frequencies



Some polar sidechains: ± charges at distance 2.2Å

Charged sidechains form salt bridge networks



Charged sidechains

Nonlocal models

Survey of results from two papers:

Xie Dexuan, Yi Jiang, Peter Brune, and L. Ridgway Scott.
A fast solver for a nonlocal dielectric continuum model.
SISC, 34(2):B107–B126, 2012.

Xie Dexuan, Yi Jiang, and L. Ridgway Scott.
Efficient algorithms for solving a nonlocal dielectric continuum model for protein in ionic solvent
SISC, to appear

Nonlocal dielectric

Linearized Poisson-Boltzmann equation:

$$-\nabla \cdot (\epsilon \nabla \Phi(\mathbf{r})) = \rho \tag{6}$$

In bulk dielectric, ϵ is a constant.

But in general it depends on the frequency of Φ, e.g.

$$\epsilon f = f + K * f \tag{7}$$

where

$$\hat{K}(\xi) = \frac{1}{1 + |\xi|^2}. \tag{8}$$

So convolution with K is the inverse of a PDE.

Can solve using a system of two PDEs.

Free energy differences

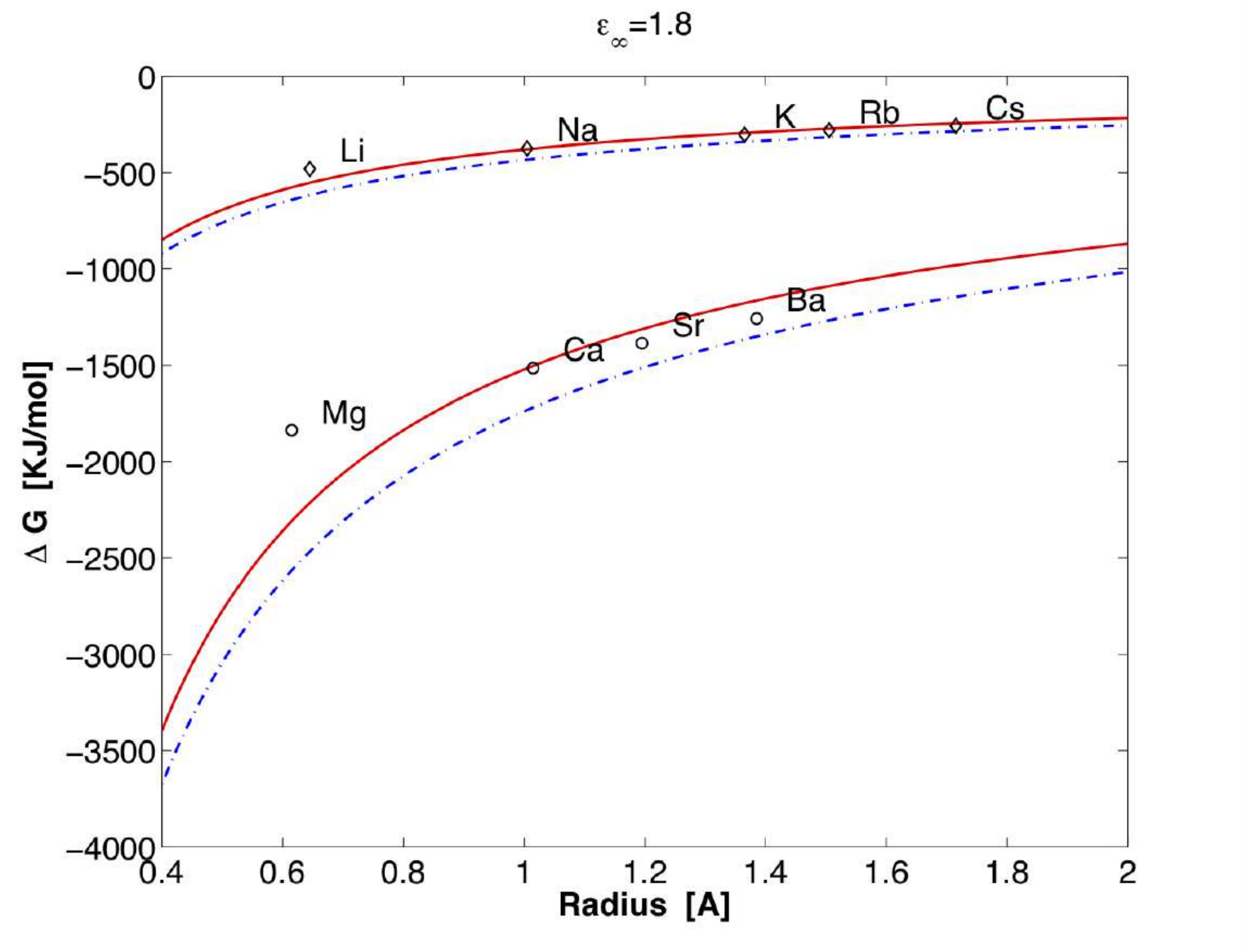


Figure 3: Comparisons of analytical free energy differences calculated from the nonlocal dielectric model with two values of λ ($\lambda = 15$Å and $\lambda = 30$Å) and the values from chemical experiments.

Singularity resolution

Step 0: Let G be the solution to $\Delta G = \rho$ in all space:

$$G(x) = \sum \frac{q_i}{|x - x_i|}.$$

Find $u_0 \in H_0^1(\Omega)$ such that

$$a_0(u_0, w) = \ell_0(w) \qquad \forall w \in H_0^1(\Omega), \tag{9}$$

where ℓ_0 and a_0 are linear and bilinear forms as defined by ($\lambda > 0$ is a model parameter)

$$\begin{aligned} a_0(u_0, w) &= \lambda^2 \int_\Omega \nabla u_0 \cdot \nabla w \, d\mathbf{r} + \int_\Omega u_0 w \, d\mathbf{r}, \\ \ell_0(w) &= \int_\Omega G(\mathbf{r}) w(\mathbf{r}) \, d\mathbf{r}. \end{aligned} \tag{10}$$

Nonlocal model with protein

For $\underset{\sim}{\phi} = (\Phi, u)$ and $\underset{\sim}{v} = (v_1, v_2)$, define

$$\begin{aligned} a(\underset{\sim}{\phi}, \underset{\sim}{v}) = & \int_\Omega \epsilon(\mathbf{r}) \nabla \Phi(\mathbf{r}) \cdot \nabla v_1(\mathbf{r}) d\mathbf{r} \\ & + (\epsilon_s - \epsilon_\infty) \int_{D_s} \nabla u(\mathbf{r}) \cdot \nabla v_1(\mathbf{r}) d\mathbf{r} \\ & + \lambda^2 \int_\Omega \nabla u(\mathbf{r}) \cdot \nabla v_2(\mathbf{r}) d\mathbf{r} + \int_\Omega (u(\mathbf{r}) - \Phi(\mathbf{r})) v_2(\mathbf{r}) d\mathbf{r}, \end{aligned} \tag{11}$$

where ϵ_p, ϵ_s, ϵ_∞ and λ are constants, and

$$\epsilon(\mathbf{r}) = \begin{cases} \epsilon_p, & \mathbf{r} \in D_p \text{ (protein)}, \\ \epsilon_\infty, & \mathbf{r} \in D_s \text{ (solvent)}, \end{cases} \tag{12}$$

Nonlocal variational equations

Find $\underset{\sim}{\phi}_1 = (\Psi, u_1), \underset{\sim}{\phi}_2 = (\tilde{\Phi}, u_2) \in \mathcal{V}$ such that

$$\begin{aligned} a(\underset{\sim}{\phi}_1, \underset{\sim}{v}) &= \ell_1(\underset{\sim}{v}) \qquad \forall \underset{\sim}{v} \in \mathcal{V}, \\ a(\underset{\sim}{\phi}_2, \underset{\sim}{v}) &= \ell_2(\underset{\sim}{v}) \qquad \forall \underset{\sim}{v} \in \mathcal{V}, \end{aligned} \tag{13}$$

where $\ell_1(\underset{\sim}{v})$ and $\ell_2(\underset{\sim}{v})$ are two linear forms as defined by

$$\begin{aligned} \ell_1(\underset{\sim}{v}) &= (\epsilon_\infty - \epsilon_s) \int_{D_p} \nabla u_0(\mathbf{r}) \cdot \nabla v_1(\mathbf{r}) \mathrm{d}\mathbf{r} \\ &\quad + (\epsilon_p - \epsilon_\infty) \int_{D_s} \nabla G(\mathbf{r}) \cdot \nabla v_1(\mathbf{r}) \mathrm{d}\mathbf{r}, \\ \ell_2(\underset{\sim}{v}) &= \frac{1}{\epsilon_0} \sum_{i=1}^{n} q_i \int_{D_s} c_i(\mathbf{r}) v_1(\mathbf{r}) \mathrm{d}\mathbf{r}. \end{aligned} \tag{14}$$

Surface mesh of BPTI

The nonlocal model solution Φ is given by

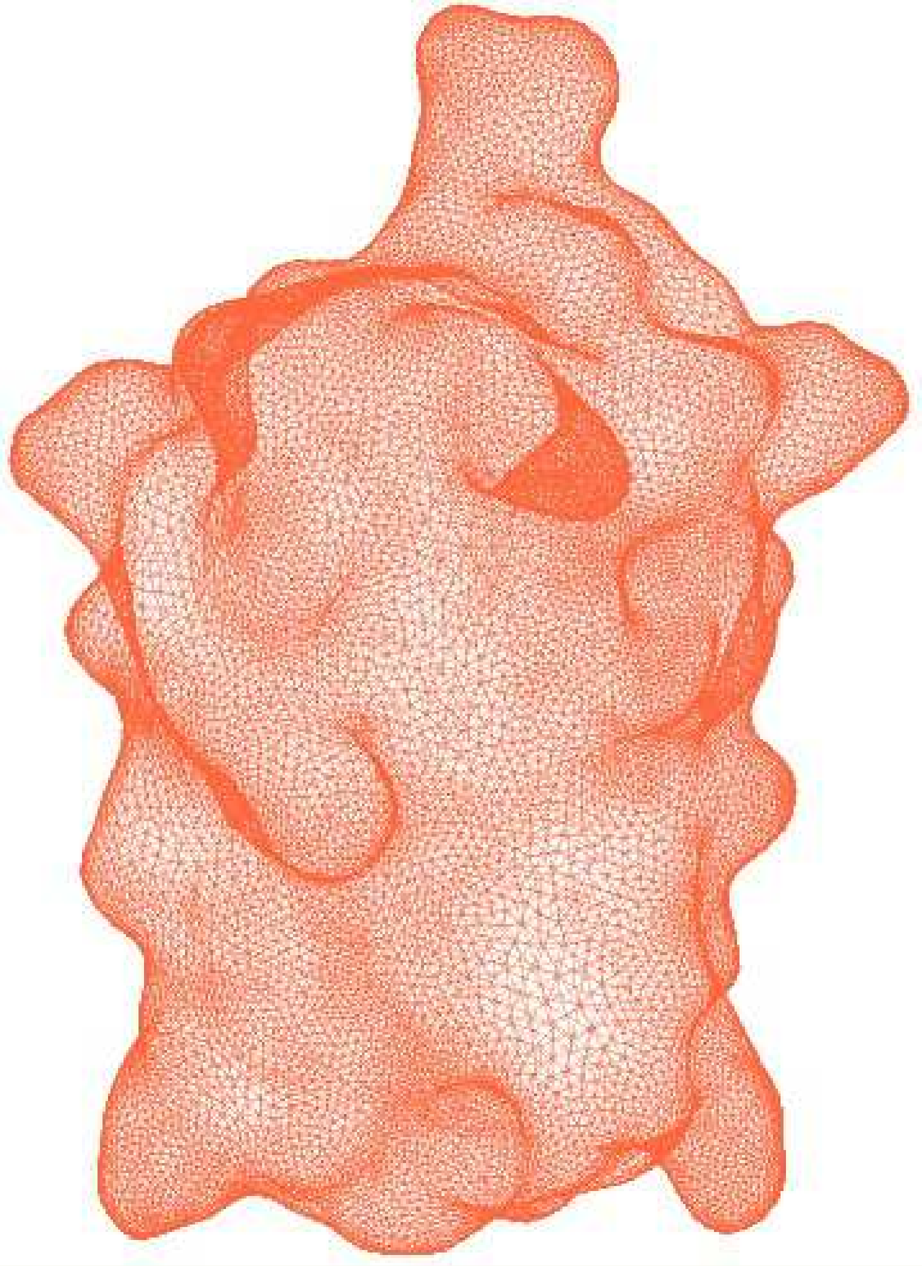
$$\Phi = \Psi + \tilde{\Psi} + G$$

Key point: singularity G is added at the end.
∇G appears in ℓ_1 only on the solvent domain D_s where there are no fixed charges.
The auxiliary variables satisfy

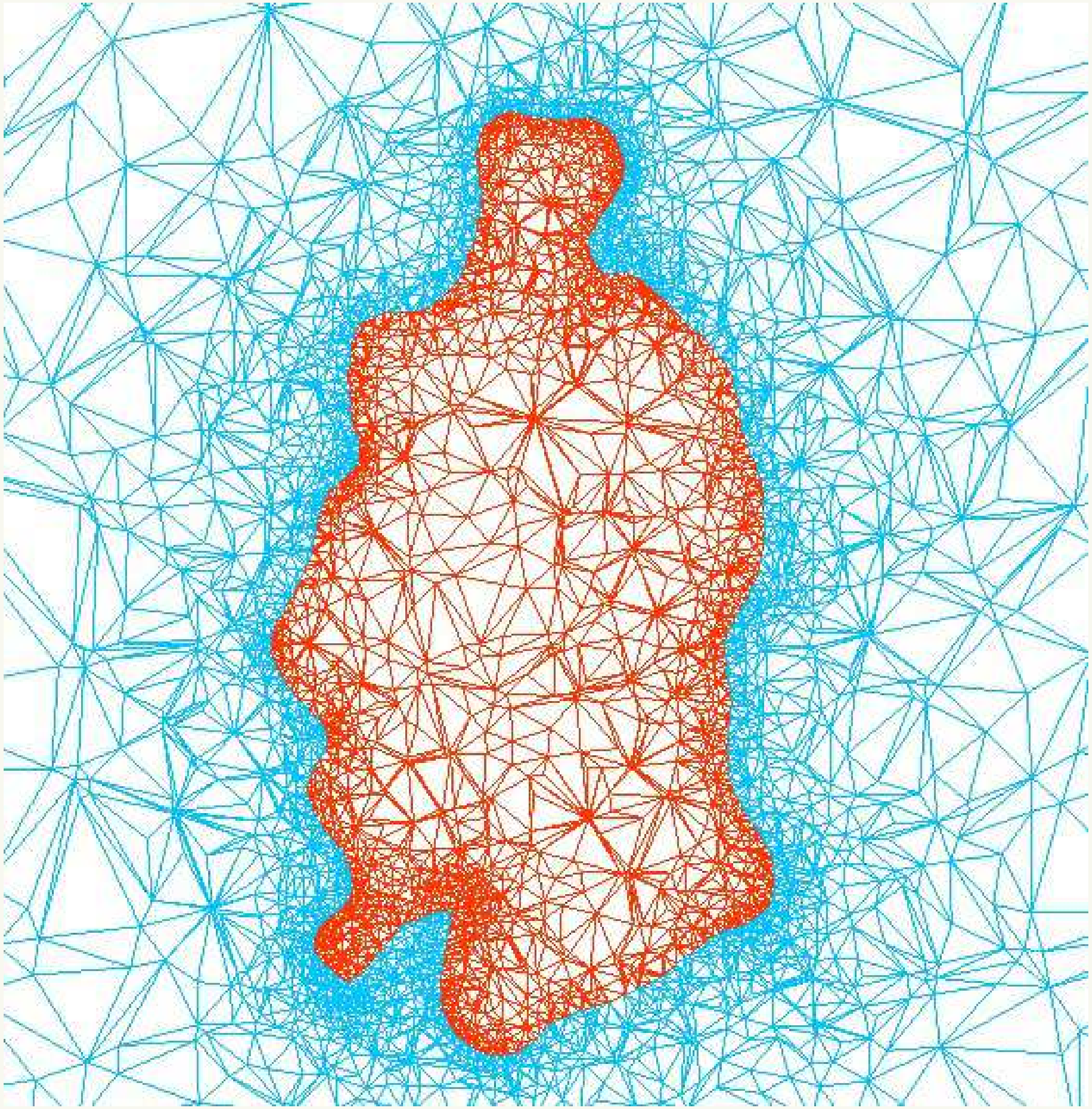
$$u = \Phi * Q_\lambda, u_1 = \Psi * Q_\lambda, u_2 = \tilde{\Psi} * Q_\lambda$$

$$u = u_0 + u_1 + u_2$$

Surface mesh of BPTI



Mesh cross-section around BPTI



Surface electrostatics of BPTI

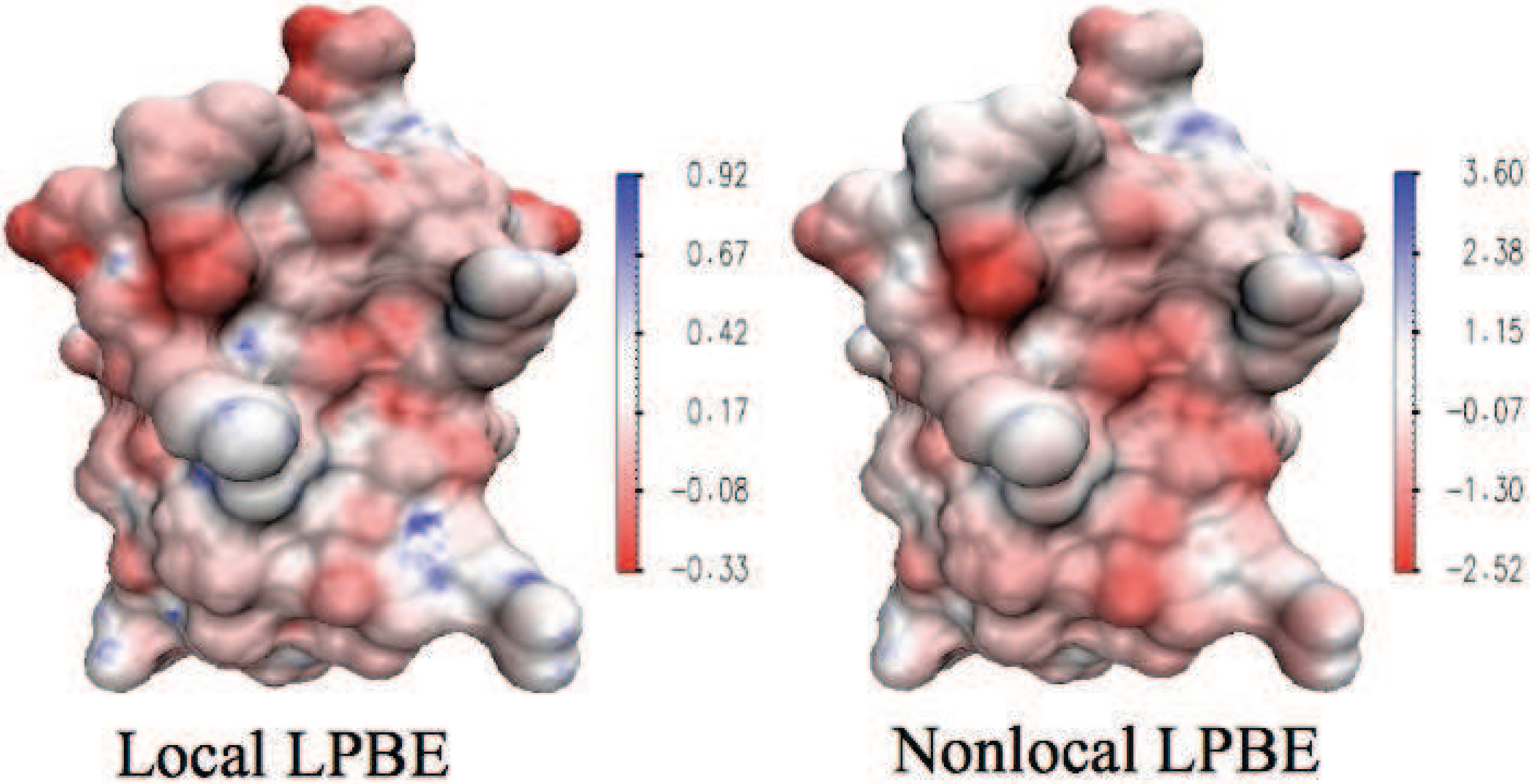


Figure 4: Nonlocal model yields different potential at protein surface.

Nonlinear models

Polarization field $\nabla\phi_\gamma$ saturates for large fixed fields:

$$\lim_{|\nabla\phi|\to\infty}(1-\varepsilon)\nabla\phi = \lim_{|\nabla\phi|\to\infty}\nabla\phi_\gamma = C, \tag{15}$$

One simple model that satisfies (15) is

$$\varepsilon(x) = \varepsilon_0 + \frac{\varepsilon_1}{1+\lambda|\nabla\phi(x)|} \tag{16}$$

for some constants ε_0, ε_1, and λ.

Both the nonlocal and nonlinear models of the dielectric response have the effect of representing frequency dependence of the dielectric effect.

$|\nabla\phi(x)|$ provides a proxy for frequency content, although it will not reflect accurately high-frequency, low-power electric fields.

Combination of nonlocal and nonlinear dielectric models may be needed.

Local model for dielectric effect?

Hydrophobic (CH_n) groups remove water locally.

This causes a reduction in ε locally.
(Resulting increase in ϕ makes dehydrons sticky.)

This can be quantified and used to predict binding sites.

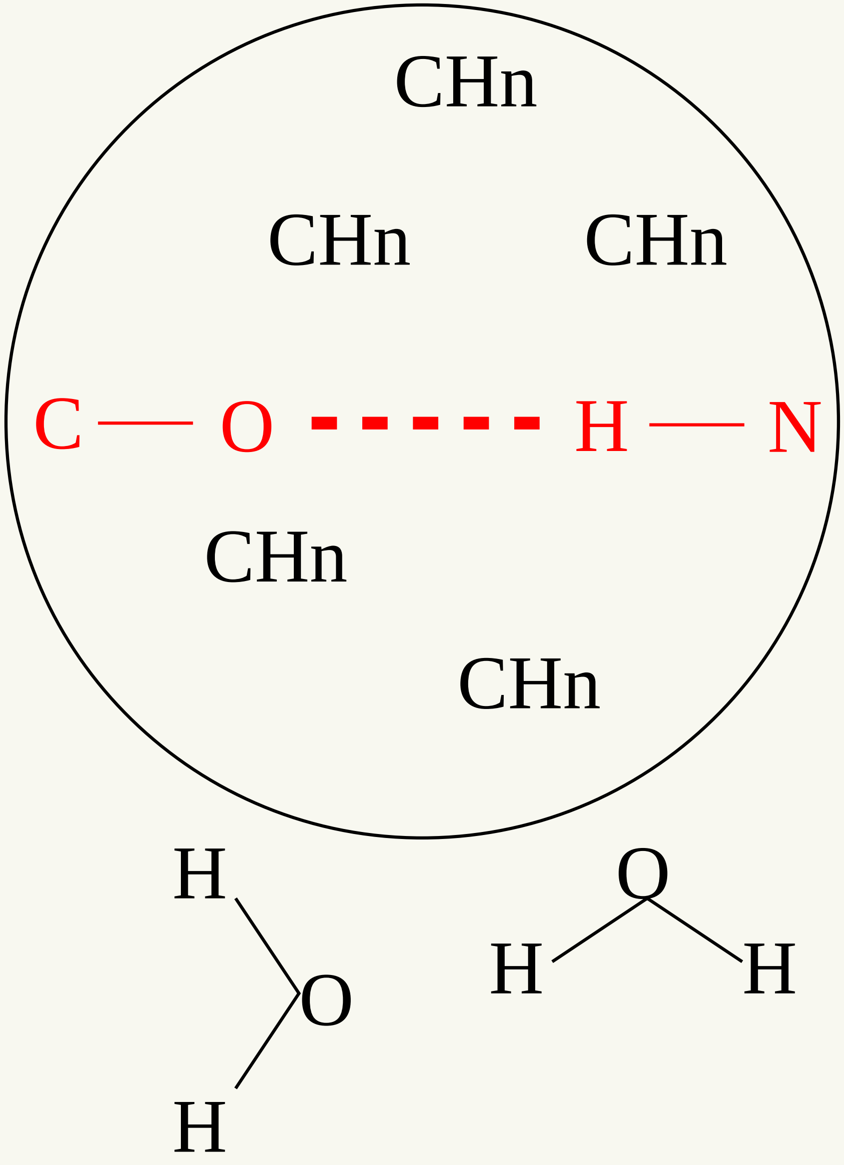
The placement of hydrophobic groups near an electrostatic bond is called wrapping.

Like putting insulation on an electrical wire.

(Wrapping modifies dielectric effect)

We can see this effect on a single hydrogen bond.

Wrapping protects hydrogen bond from water



CHn

CHn CHn

C — O - - - - - H — N

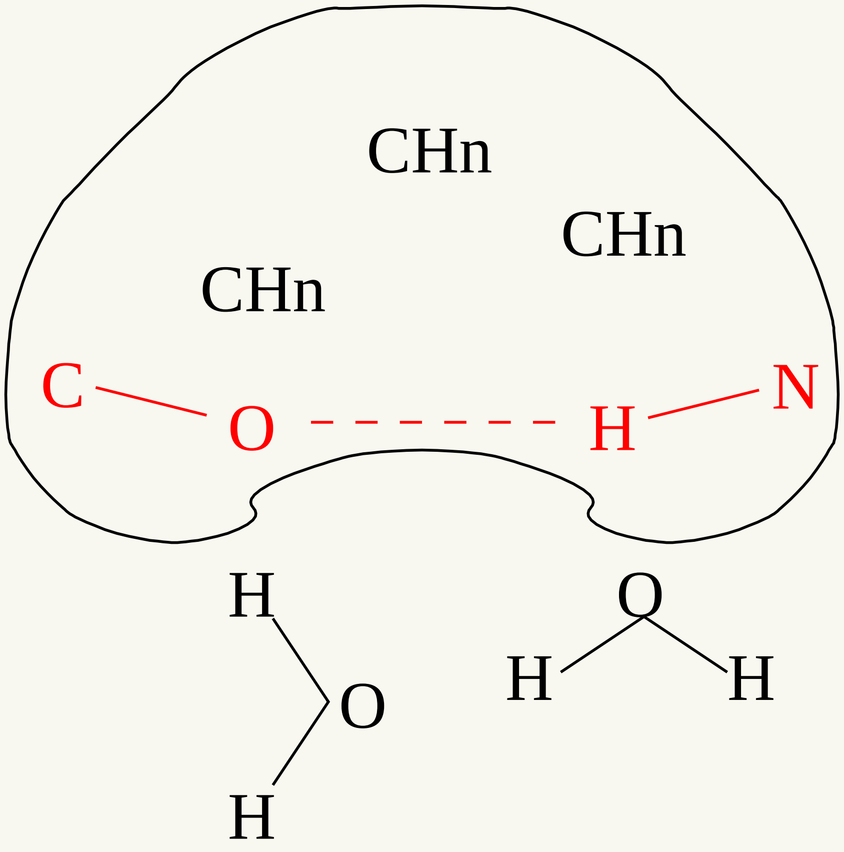
CHn

CHn

H O H

H O H

Well wrapped hydrogen bond



CHn

CHn

CHn

C — O - - - - - - H — N

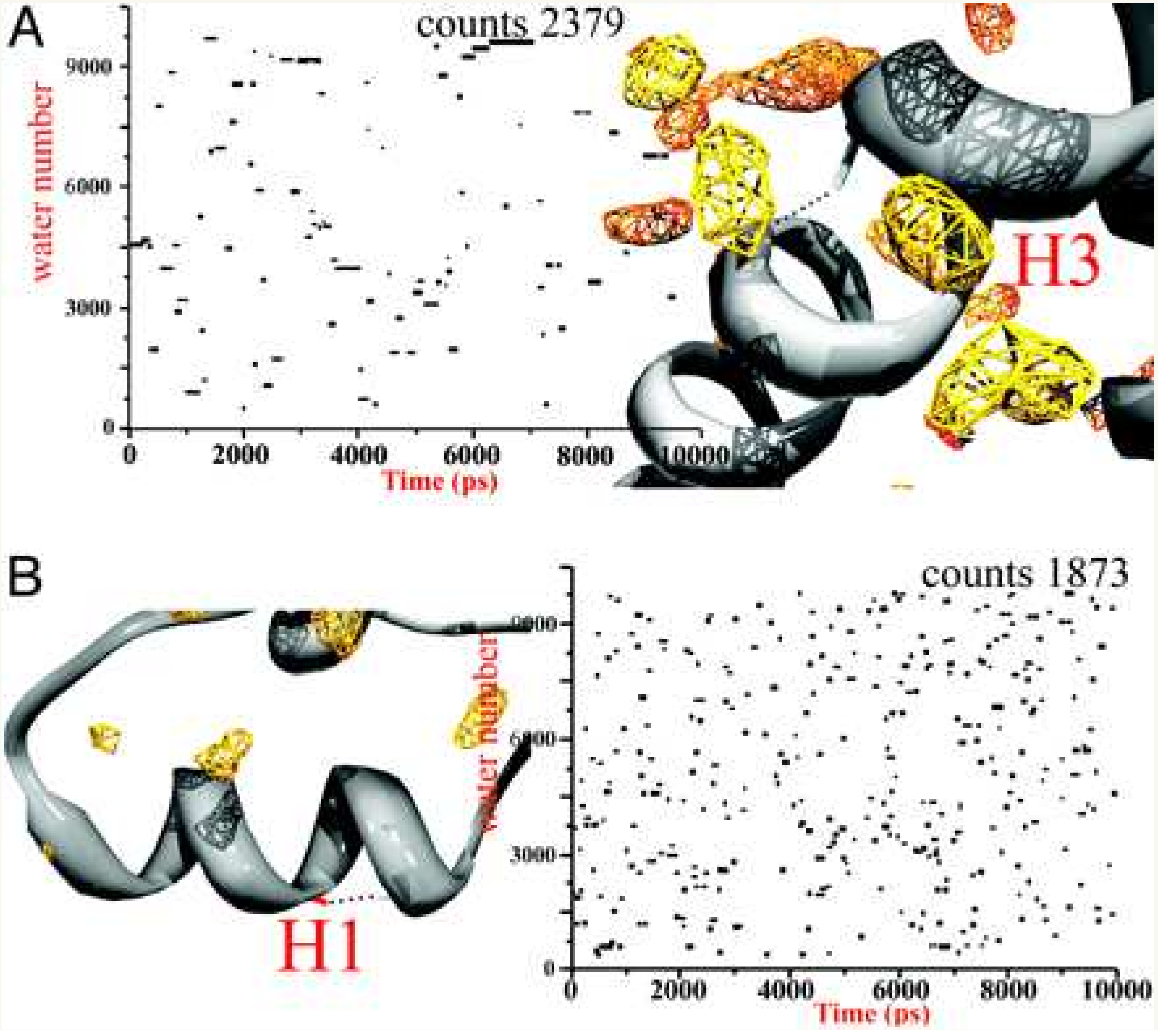
H O H

H O H

Underwrapped hydrogen bond

Extent of wrapping changes nature of hydrogen bond

Hydrogen bonds (B) that are not protected from water do not persist.



From De Simone, et al., PNAS 102 no 21 7535-7540 (2005)

Dynamics of hydrogen bonds and wrapping

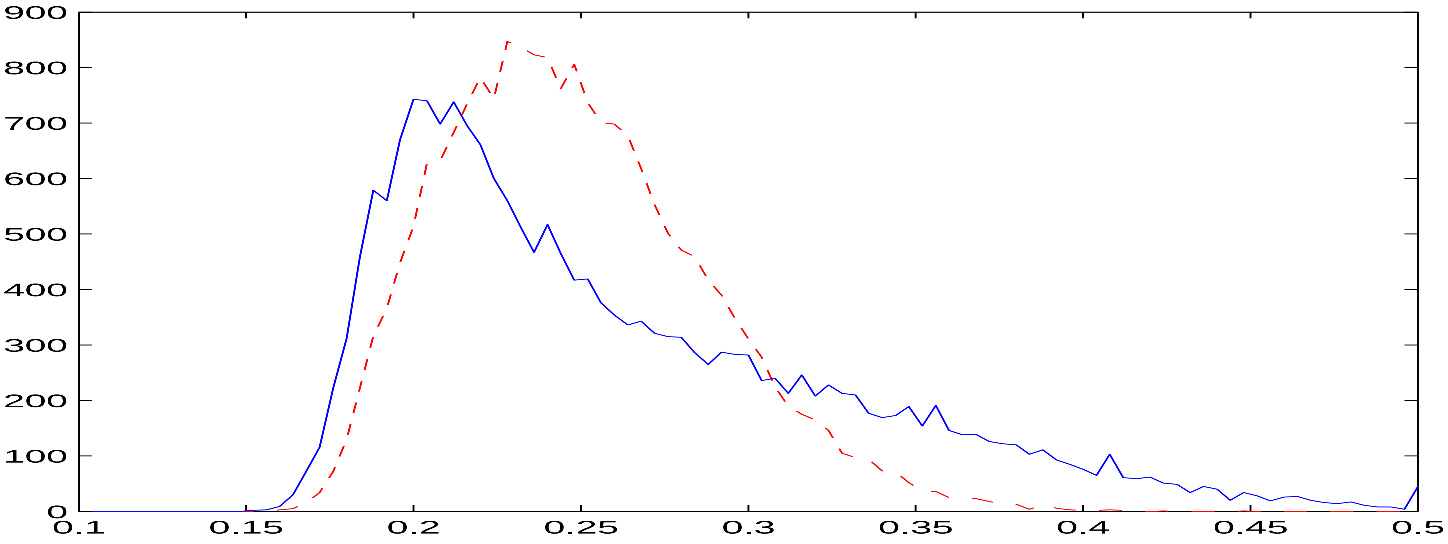
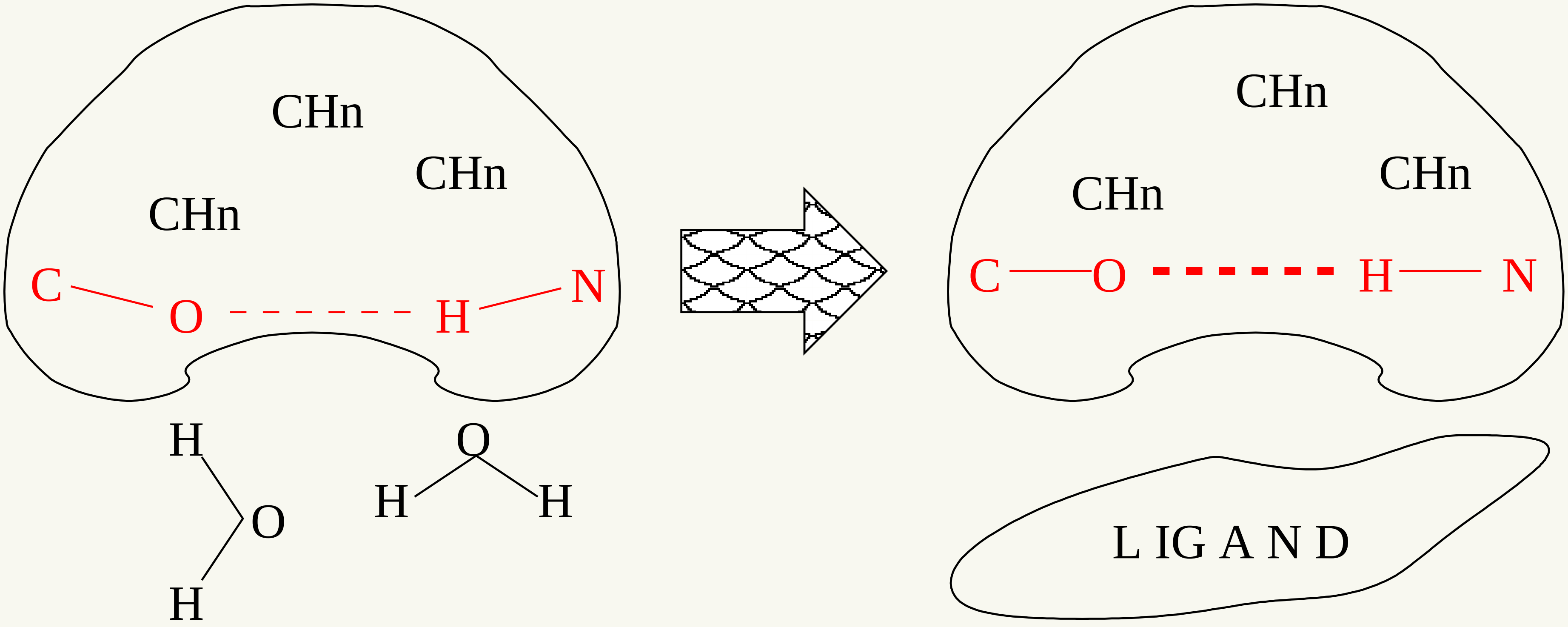


Figure 5: Distribution of bond lengths for two hydrogen bonds formed in a structure of the sheep prion [2]. Horizontal axis measured in nanometers, vertical axis represents numbers of occurrences taken from a simulation with $20,000$ data points with bin widths of 0.1 Ångstrom. Distribution for the well-wrapped hydrogen bond (H3) has smaller mean value but a longer (exponential) tail, whereas distribution for the underwrapped hydrogen bond (H1) has larger mean but Gaussian tail.

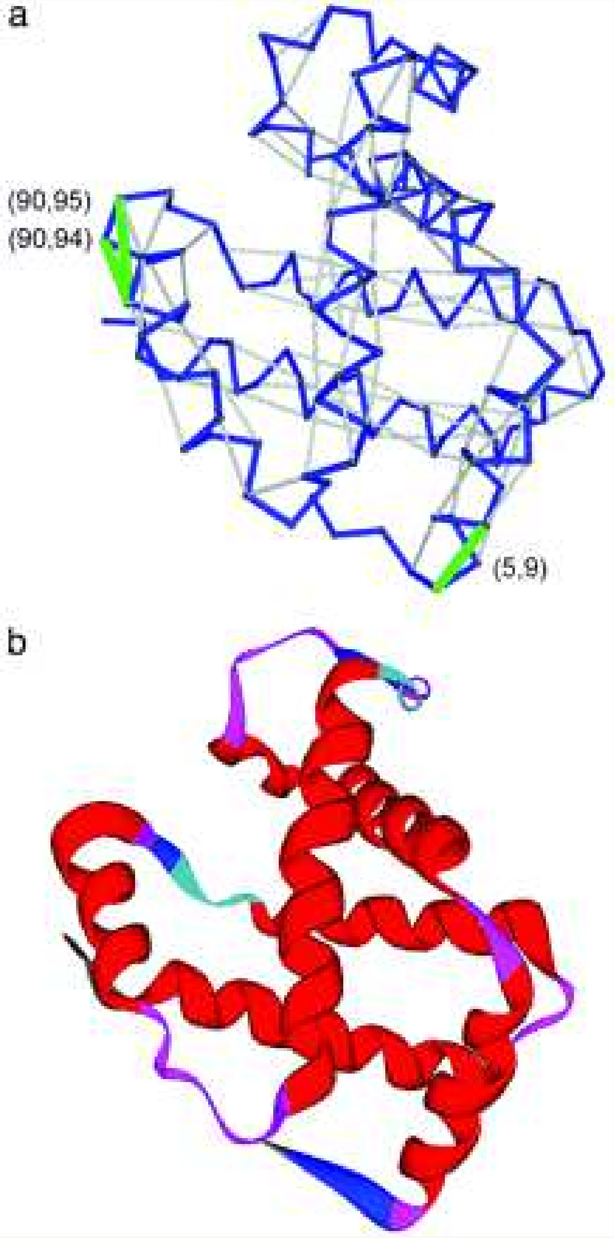
Ligand binding removes water



Binding of ligand changes underprotected hydrogen bond (high dielectric) to strong bond (low dielectric)

No intermolecular bonds needed!

Dehydrons



Dehydrons

in human hemoglobin, From
PNAS 100: 6446-6451 (2003) Ariel Fernandez, Jozsef Kardos, L. Ridgway Scott, Yuji Goto, and R. Stephen Berry. Structural defects and the diagnosis of amyloidogenic propensity.

Well-wrapped hydrogen bonds are grey, and dehydrons are green.

The standard ribbon model of “structure” lacks indicators of electronic environment.

Mathematical explanation

Charges ρ induce an electric field $e = \nabla\phi$ given by

$$\nabla \cdot (\epsilon\nabla\phi) = \nabla \cdot (\epsilon e) = \rho$$

Energy $= -\int \rho\phi\, dx$

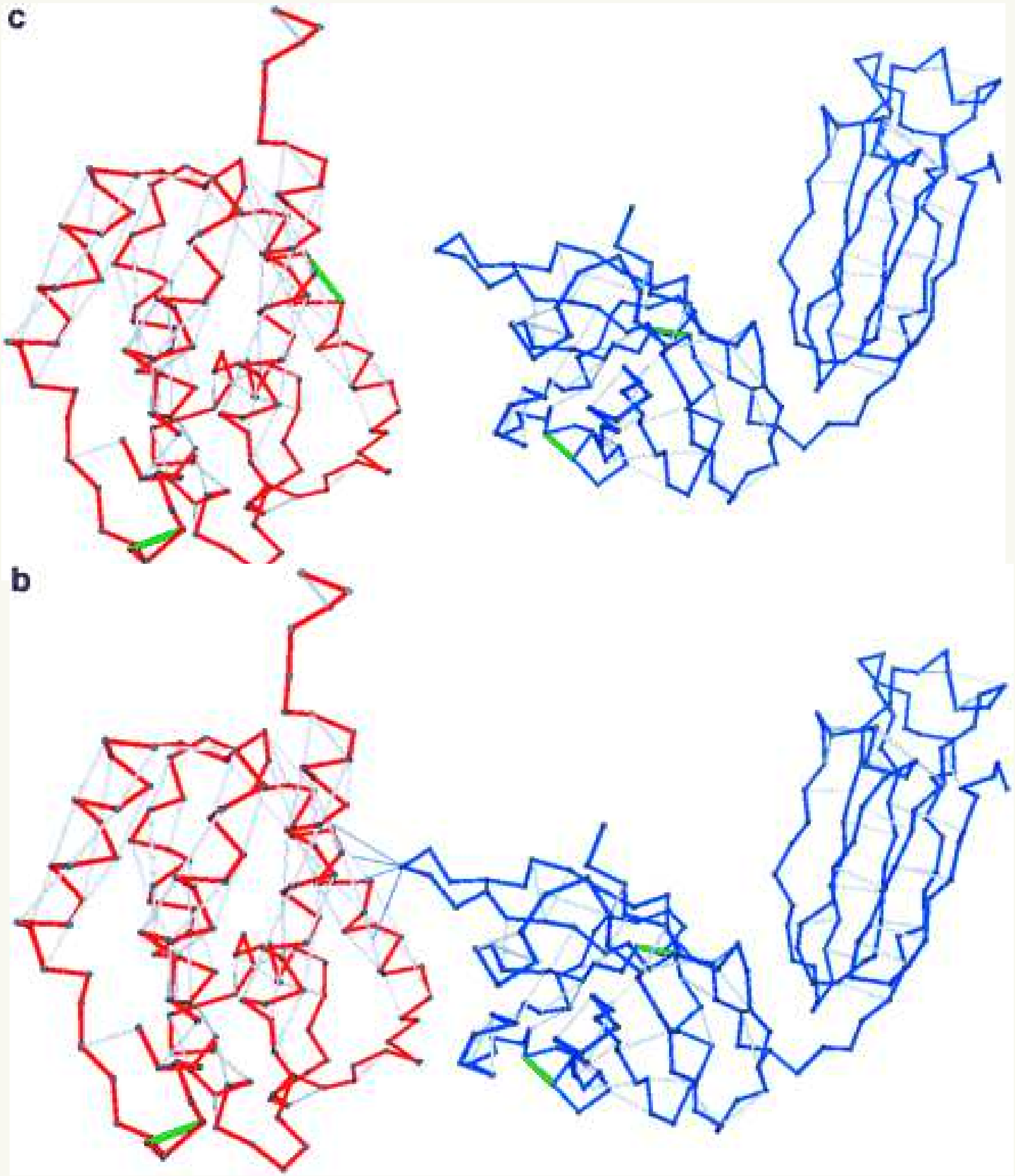
Hydrophobicity affects the operator ϵ: removing water reduces ϵ.

When ϵ goes down, ϕ goes up.

Hydrophilic groups contribute to the right-hand side ρ.

Hydrophobicity and hydrophilic are orthogonal, not opposites.

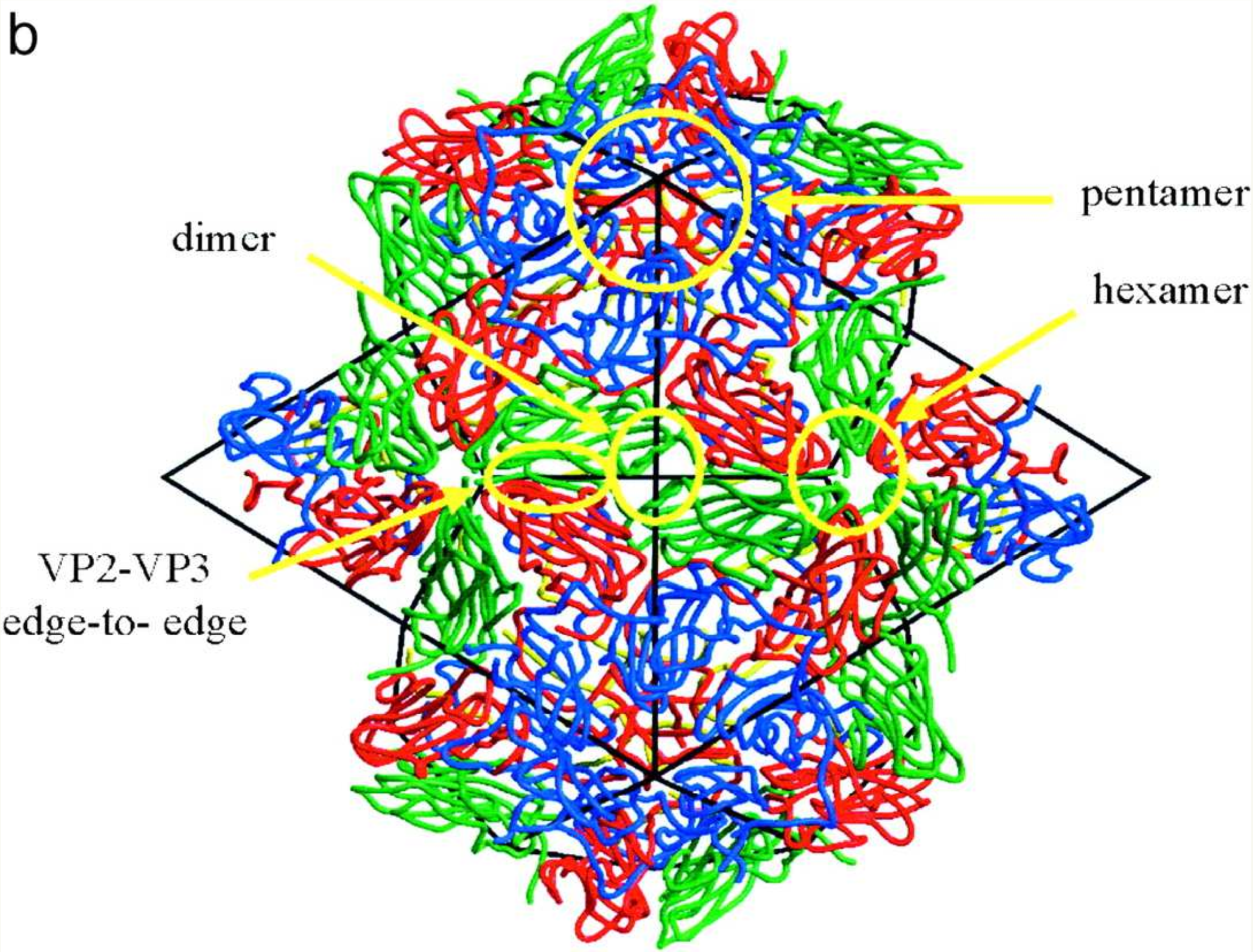
Antibody binding to HIV protease



The HIV protease has a dehydron at an antibody binding site.

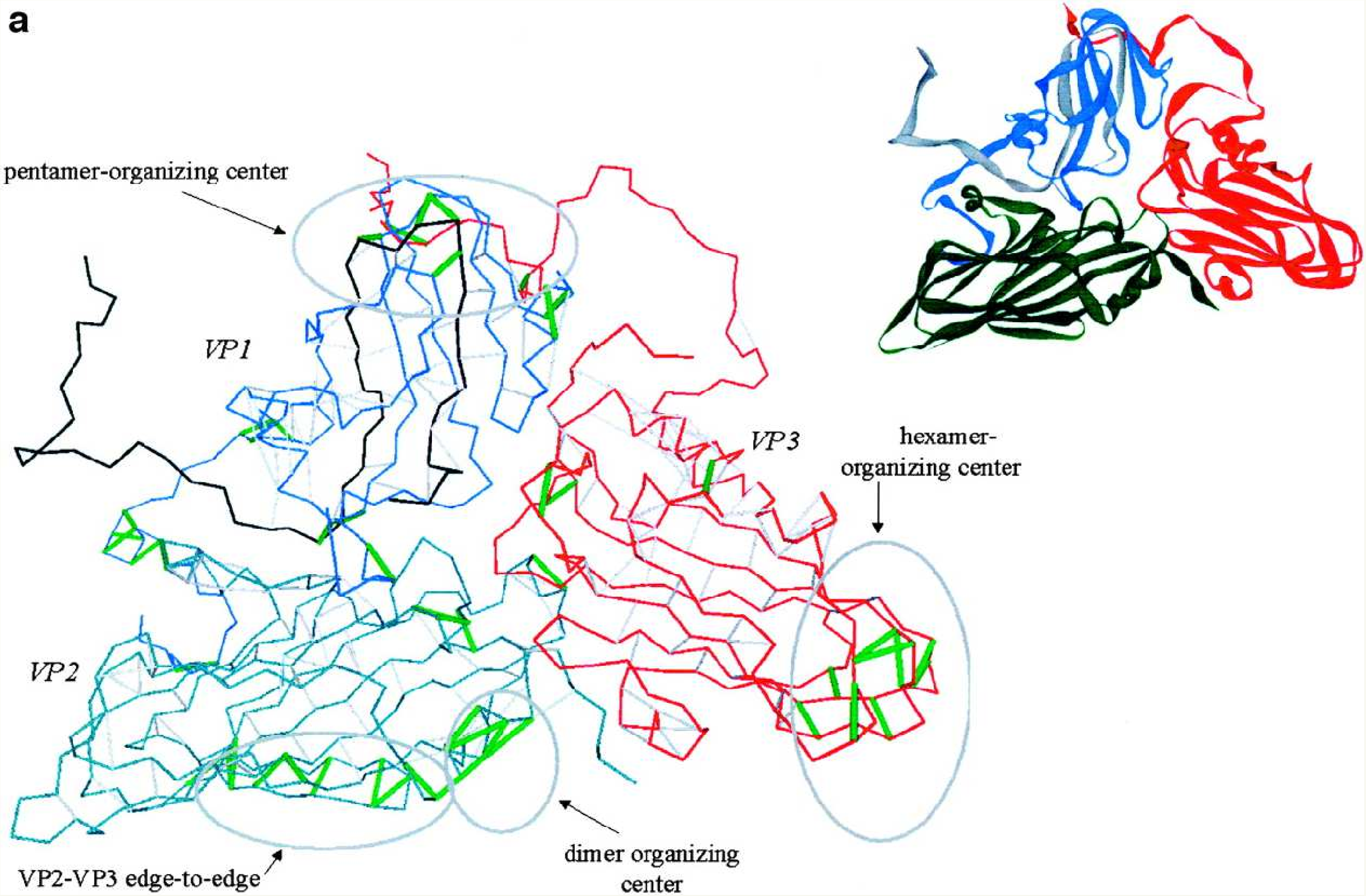
When the antibody binds at the dehydron, it wraps it with hydrophobic groups.

A model for protein-protein interaction



Foot-and-mouth disease virus assembly from small proteins.

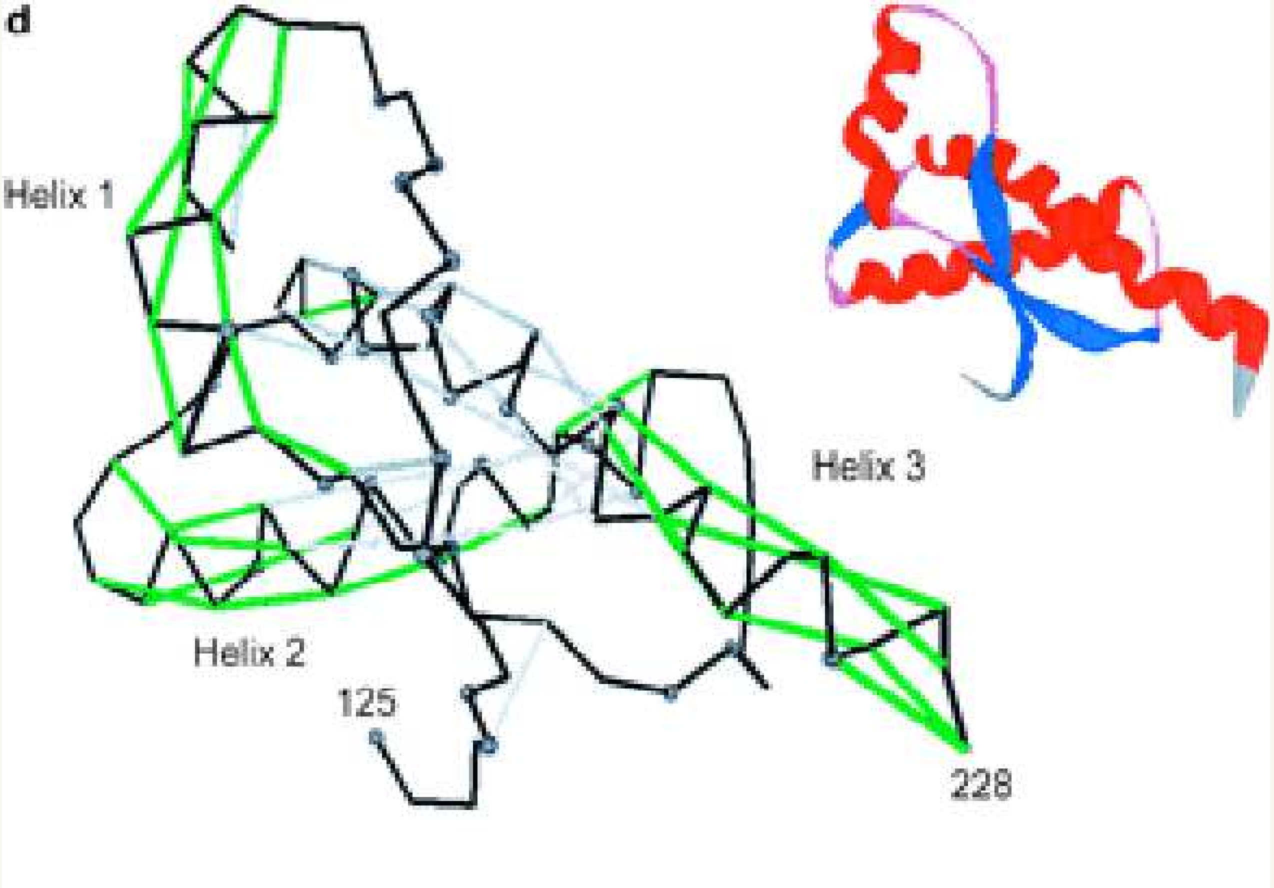
Dehydrons guide binding



Dehydrons guide binding of component proteins VP1, VP2 and VP3 of foot-and-mouth disease virus.

Extreme interaction: amyloid formation

Standard application of bioinformatics: look at distribution tails.
If some is good, more may be better, but too many may be bad.
Too many dehydrons signals trouble: the human prion.



From PNAS 100: 6446-6451 (2003) Ariel Fernandez, Jozsef Kardos, L. Ridgway Scott, Yuji Goto, and R. Stephen Berry. Structural defects and the diagnosis of amyloidogenic propensity.

Genetic code

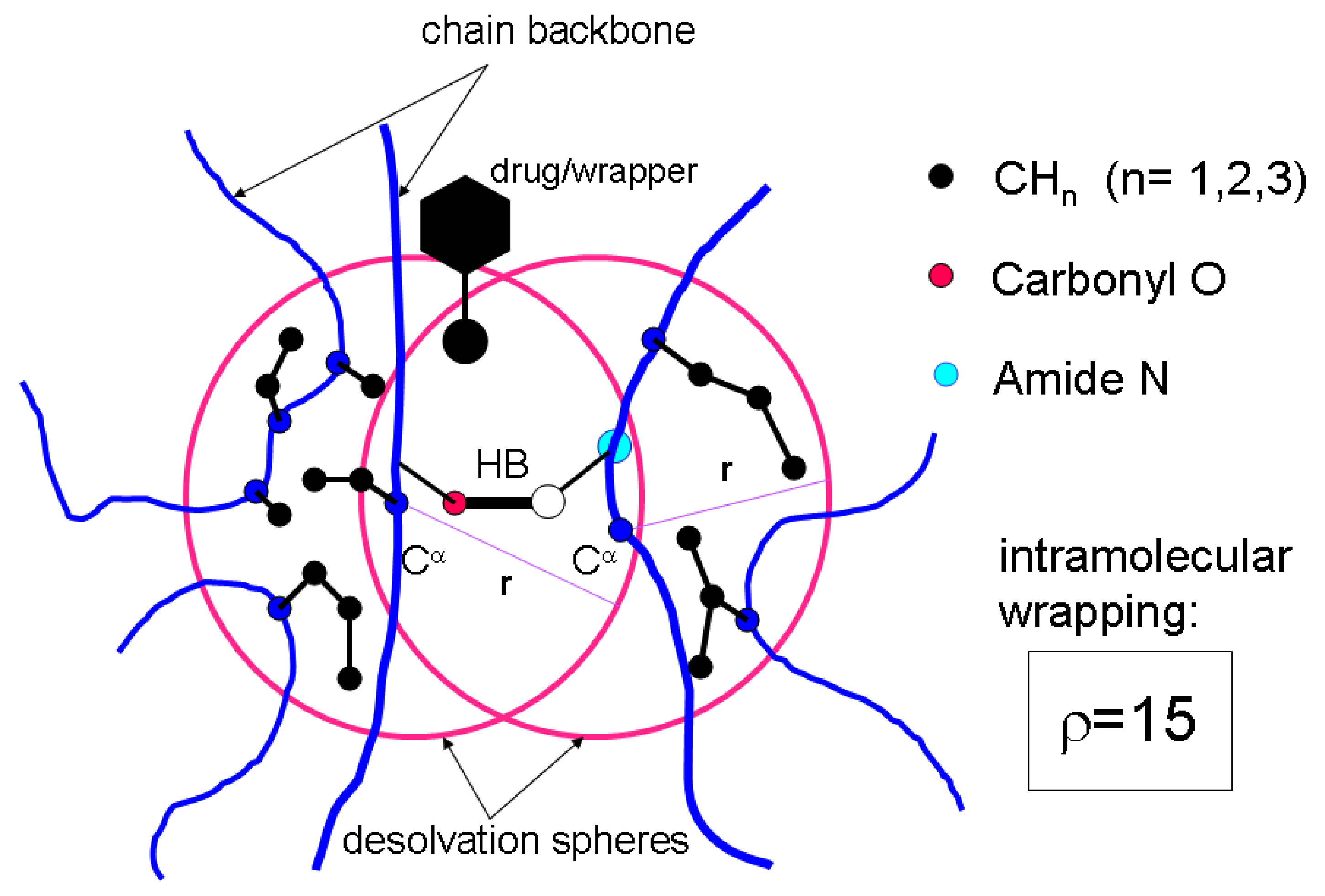
Genetic code minimizes changes of polarity due to single-letter codon mutations, but it facilitates changes in wrapping due to single-letter codon mutations.

	Second Position								
First Position	u		c		a		g		Third Position
u	uuu	Phe 7 | |	ucu	Ser 0 + −	uau	Tyr 6 + −	ugu	Cys 0 + −	u
	uuc		ucc		uac		ugc		c
	uua	Leu 4 | |	uca		uaa	stop	uga	stop	a
	uug		ucg		uag	stop	ugg	Trp 7 + −	g
c	cuu	Leu 4 | |	ccu	Pro 2 | |	cau	His 1 + −	cgu	Arg 2 + +	u
	cuc		ccc		cac		cgc		c
	cua		cca		caa	Gln 2 + −	cga		a
	cug		ccg		cag		cgg		g
a	auu	Ile 4 | |	acu	Thr 1 + −	aau	Asn 1 + −	agu	Ser 0 + −	u
	auc		acc		aac		agc		c
	aua		aca		aaa	Lys 3 + +	aga	Arg 2 + +	a
	aug	Met 1 + −	acg		aag		agg		g
g	guu	Val 3 | |	gcu	Ala 1 | |	gau	Asp 1 − −	ggu	Gly 0 + −	u
	guc		gcc		gac		ggc		c
	gua		gca		gaa	Glu 2 − −	gga		a
	gug		gcg		gag		ggg		g

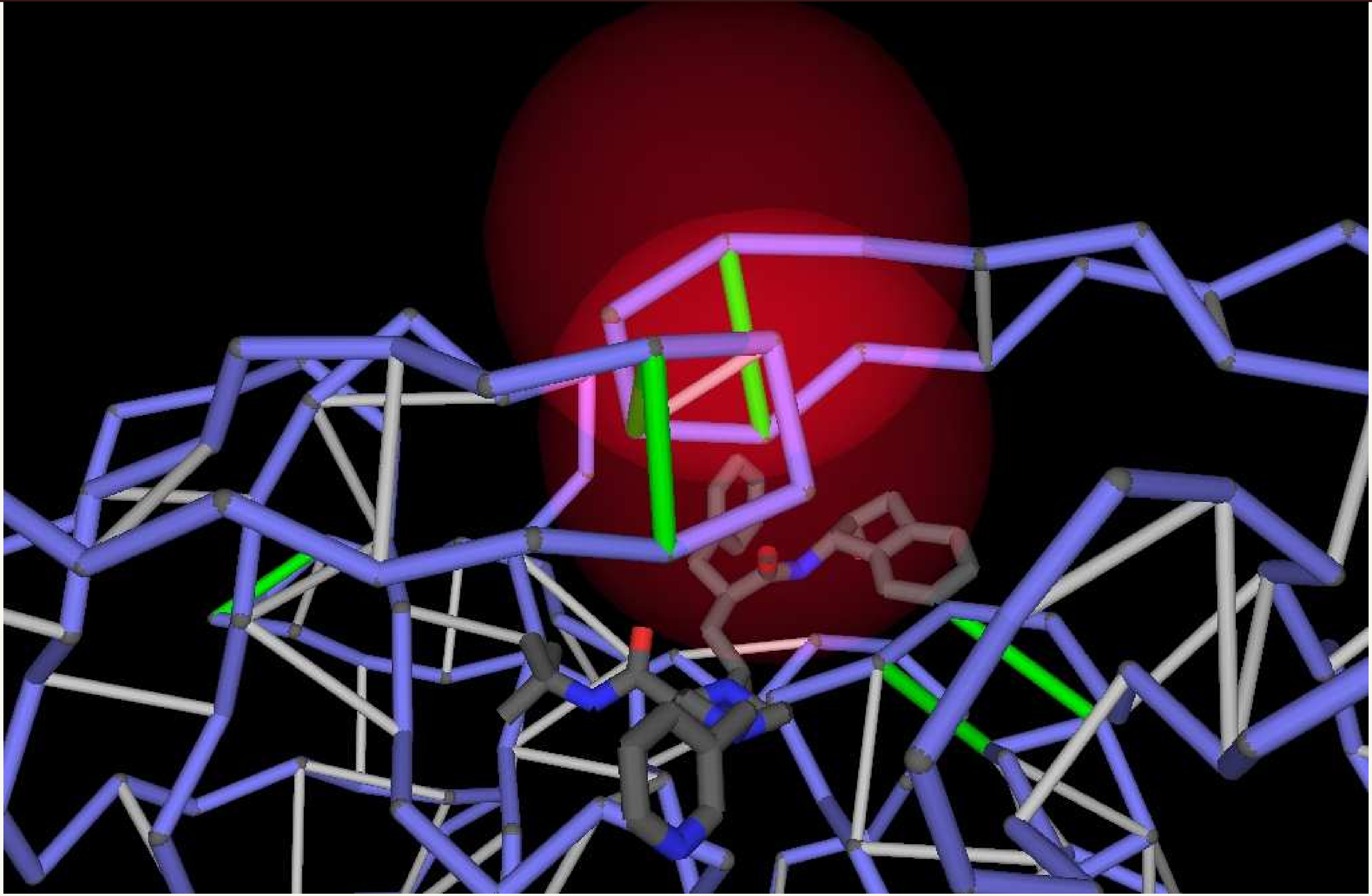
First digit after residue name is amount of wrapping. Second indicator is polarity; | |: nonpolar, + −: polar, − −: negatively charged, + +: positively charged.

Drug ligand wrapping

Drug ligand provides additional non-polar carbonaceous group(s) in the desolvation domain, enhancing the wrapping of a hydrogen bond.



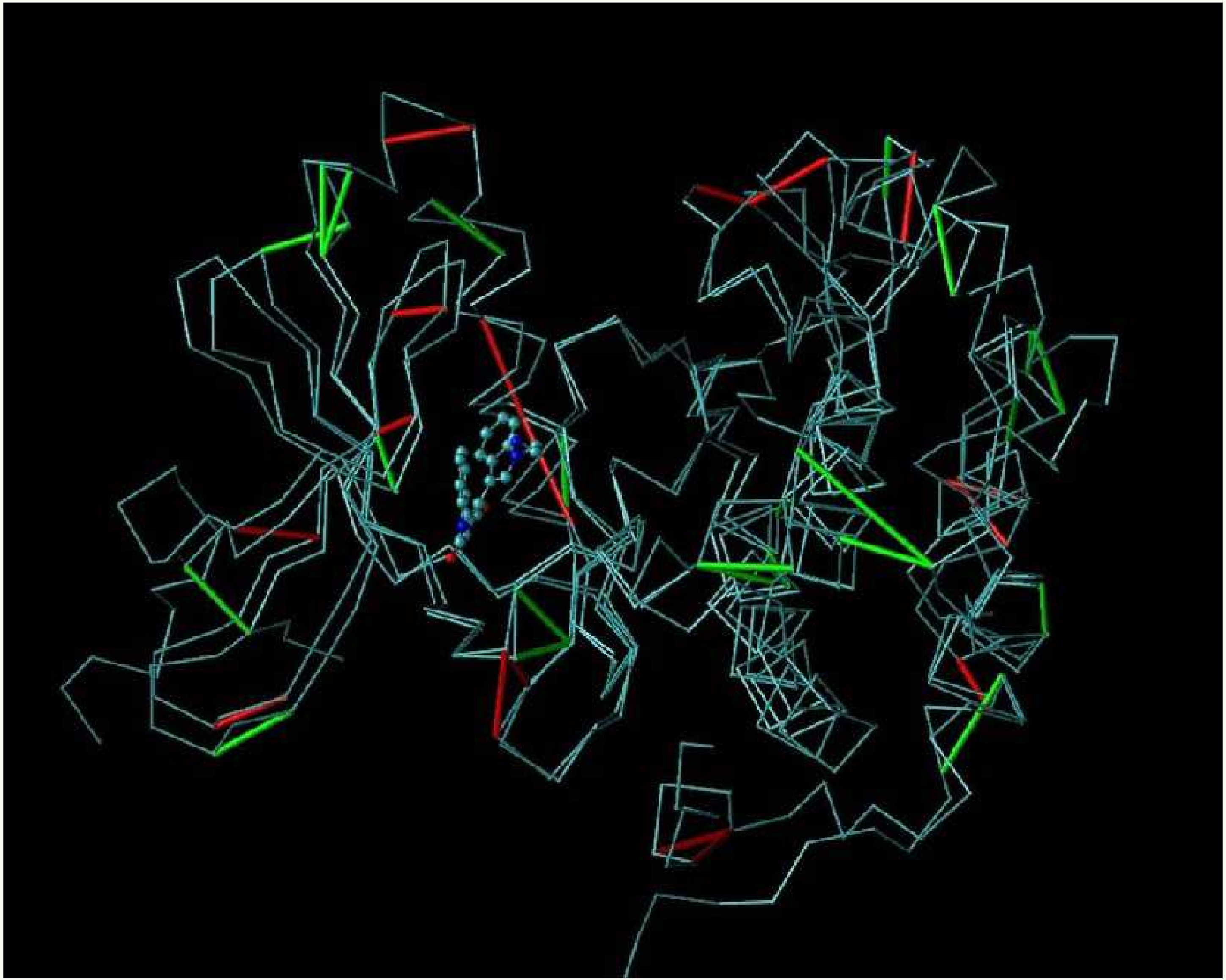
Desolvation spheres



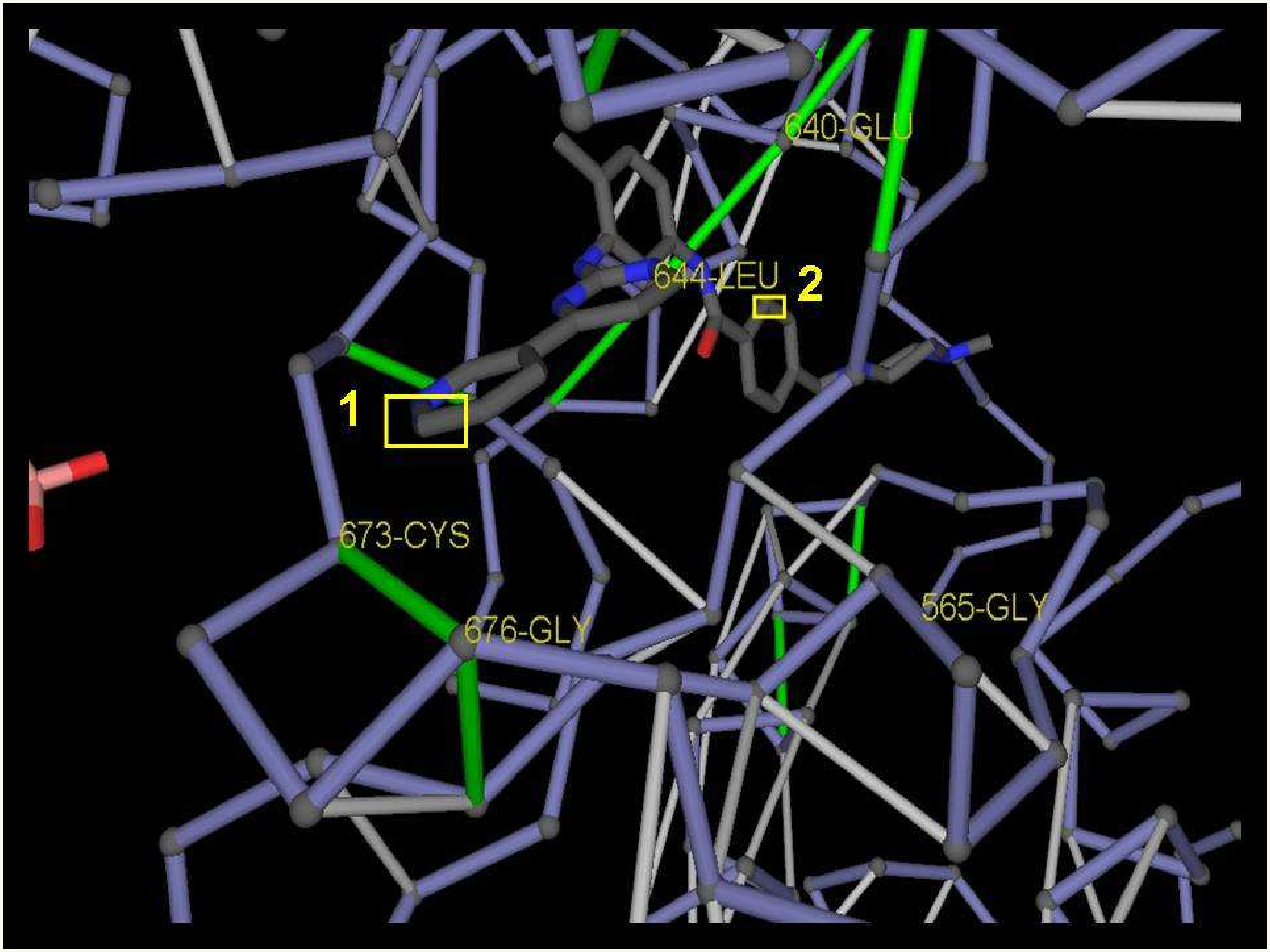
Desolvation spheres for flap Gly-49–Gly-52 dehydron containing nonpolar groups of the wrapping inhibitor.

Aligned paralogs

Aligned backbones for two paralog kinases; dehydrons for Chk1 are marked in green and those for Pdk1 are in red.

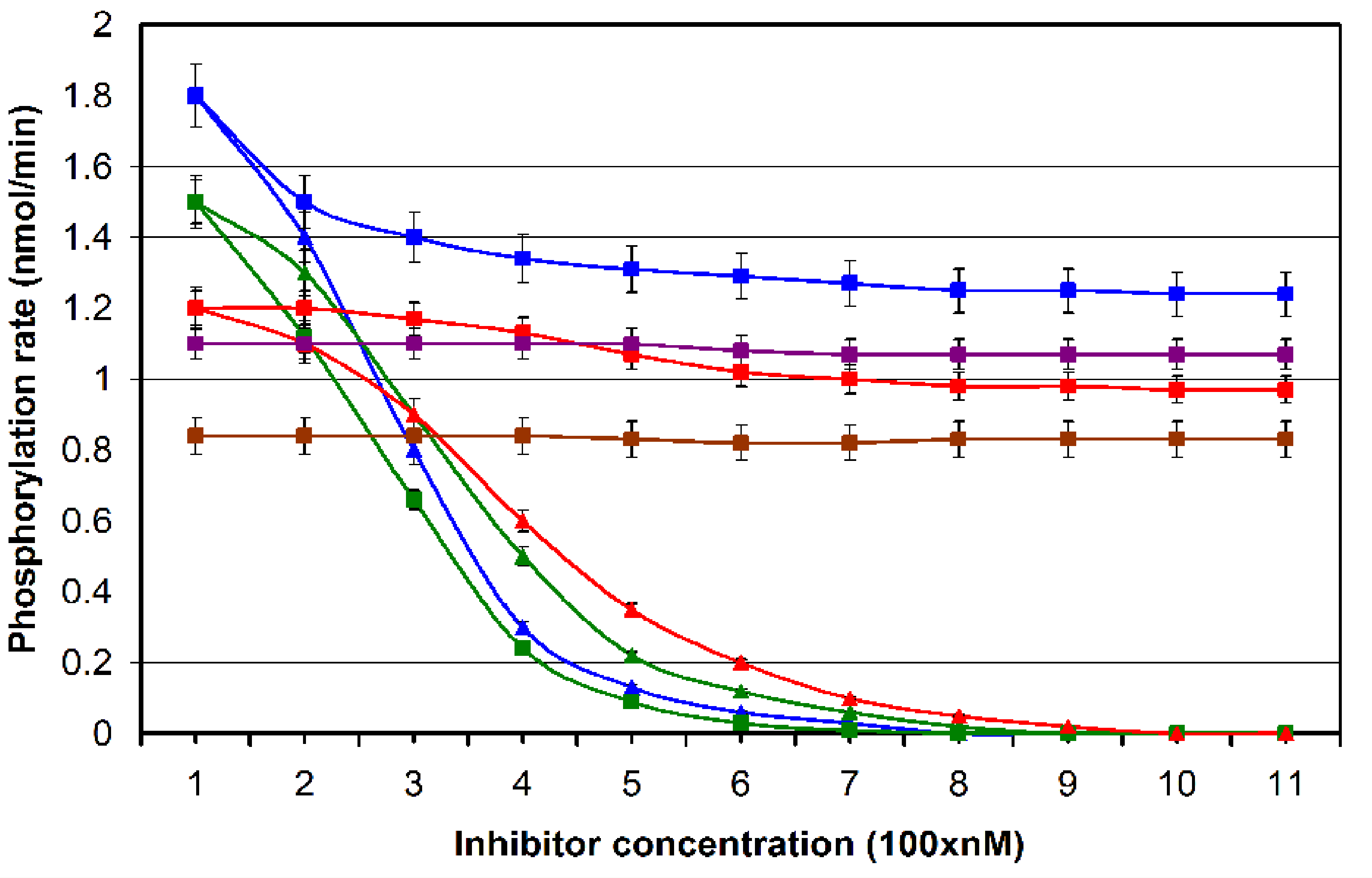


Selective wrapper



Dehydron Cys673-Gly676 in C-Kit is not conserved in its paralogs Bcr-Abl, Lck, Chk1 and Pdk1. By methylating Gleevec at the para position (1), the inhibitor becomes a selective wrapper of the packing defect in C-Kit.

Experimental confirmation



Phosphorylation rates from spectrophotometric assay on the five kinases Bcr-Abl (blue), C-Kit (green), Lck (red), Chk1 (purple), and Pdk1 (brown) with Gleevec (triangles) and modified Gleevec methylated at positions (1) and (2) (squares). Notice the selective and enhanced inhibition of C-Kit.

Conclusions

Some advances in solvation modeling

- Now tractable to compute nonlocal dielectric models
- Simple models (wrapping/dehydrons) give consistent predictions at picoscale
- Have been used to aid drug design

Thanks

This talk was based on joint work with

- Peter Brune (Argonne Nat. Lab.),
  Yi Jiang and Dexuan Xie (UWisconsin-Milwaukee)
- Steve Berry (U. Chicago),
  Ariel Ferndandez (Calderon Inst., Argentina),
  Chris Fraser (Bioanalytical Computing),
  Kristina Rogale Plazonic (Princeton),
  Harold Scheraga (Cornell)

We thank the Institute for Biophysical Dynamics at the University of Chicago and NSF for support.

We are also grateful to the developers of the PDB, Viper, DIP, and other biological data bases.

References

[1] P. A. Bopp, A. A. Kornyshev, and G. Sutmann. Static nonlocal dielectric function of liquid water. *Physical Review Letters*, 76:1280–1283, 1996.

[2] Alfonso De Simone, Guy G. Dodson, Chandra S. Verma, Adriana Zagari, and Franca Fraternali. Prion and water: Tight and dynamical hydration sites have a key role in structural stability. *Proceedings of the National Academy of Sciences, USA*, 102:7535–7540, 2005.

[3] P. Debye. *Polar Molecules*. Dover, New York, 1945.

[4] John Barrett Hasted. *Aqueous Dielectrics*. Chapman and Hall, 1974.

[5] U. Kaatze, R. Behrends, and R. Pottel. Hydrogen network fluctuations and dielectric spectrometry of liquids. *Journal of Non-Crystalline Solids*, 305(1):19–28, 2002.

[6] L. Ridgway Scott, Mercedes Boland, Kristina Rogale, and Ariel Fernández. Continuum equations for dielectric response to macro-molecular assemblies at the nano scale. *Journal of Physics A: Math. Gen.*, 37:9791–9803, 2004.