Accurate Calculation of Hydration Free Energies Using Macroscopic Solvent Models

Barry Honig, et al.

Presented by Maruti Sharma
AMS 691
Professor Rizzo

Outline

- Introduction to Electrostatics
- Poisson Boltzmann Equation (PBE)
- Applications of PBE
- PBE and Solvation of Free Energies
- Results
- Discussion

Classic Electrostatics

- Electrostatic interactions play key role in biological processes
- Interactions of molecules in aqueous solution
- Solute and Solvent molecules described as Force Fields
- van der Waals Forces and electrostatic terms account for most of the intermolecular interactions
- Important in Drug Design and Protein Folding
- PBE is key

Poisson Boltzmann Equation

$$\nabla \bullet [\varepsilon(r)\nabla \bullet \varphi(r)] - \varepsilon(r)\kappa(r)^2 \sinh[\varphi(r)]$$
$$+ 4\pi\rho^f / kT = 0$$

- k is Boltzmann constant
- T is temperature
- φ is the electrostatic potential
- ρ^f is the fixed charge density
- ε is dielectric constant
- ε, κ, ρ, φ are all functions of position vector r

Applications of PBE

- Description of Electric Potentials
- Diffusion Limited Processes
- pH dependent properties of proteins
- Ionic strength-dependent phenomena
- Solvation free energies of organic molecules

Calculating Electric Potentials

- Molecular Surface
- Interior regions $\varepsilon \sim (2\text{-}4)$
- Exterior regions ε of water~80
- Charge distribution represented as point charges of nuclei
- PBE is used to describe molecular surfaces

Patterns of $\varphi(r)$ in Proteins

- Applications used in visual representation of $\varphi(r)$ for proteins
- $\varphi(r)$ is dependent on specific location of charged and polar groups in proteins
- Proteins generate unique patterns
- Boundary of high ε regions and low ε regions
- Dielectric discontinuity produces distinctive shape

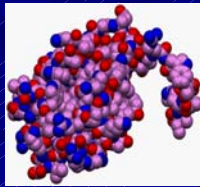## Space Model of a Protein

- Red spheres$\rightarrow$ Oxygen of negatively charged carboxylic acids
- Blue spheres$\rightarrow$ Nitrogen of basic groups
- Represents charges on active site

Calculation pK$_a$'s: pH Effects on Binding and Stability

- pK$_a$ is the proton affinity of acids and bases
- Shifts between naïve and denatured$\rightarrow$differences in pH
- Macromolecular stability effected
- Denaturation of proteins $\rightarrow$pK$_a$ shifts of amino acids
- Hydrogen bonding interactions determine pK$_a$
- PB calculations account for both the desolvation and the hydrogen bonding effects

Solvation Free Energies

- Differences in the reaction field energy of a charge distribution
- Transferred from one dielectric medium to another
- Energetics of microscopic hydrogen bonding interactions of polar molecules in the aqueous phase
- Nonpolar and electrostatic contributions included for calculating solvation free energies
- PB calculations used on solute to obtain free energy values

---

Calculation of Free Energies: Introduction

- Fast and accurate calculations of hydration free energies of small molecules
- Uses FDPB/γ method, PARSE parameters
- Examination of force field or *ab initio* parameters
- Experimental solvation energies obtained for small polar molecules
- Includes those representing amino acid side chains and polypeptide backbone
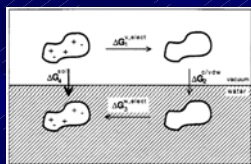

Theory and Methods: Solvation Energy Calculations

$$\Delta G_4^{sol} = \Delta G_1^{v,elect} + \Delta G_2^{c/vdw} + \Delta G_3^{w,elect}$$

- Thermodynamic cycle showing into electrostatic and van der Waals contributions to solvation energy

- $\Delta G_1^{v,\ elect}$

- $\Delta G_2^{c/vdw}$

- $\Delta G_3^{w,\ elect}$

- ΔG_4^{sol}

Theory and Methods: Electrostatic Solvation Energy Contributions

- Solute $\varepsilon = 1$
- Solvent
  - $\varepsilon = 1 \rightarrow$ vacuum
  - $\varepsilon = 80 \rightarrow$ water
- Solute and solvent are mapped onto a grid
- Electrostatic potentials calculated by Poisson Eq.
- Solvation energy given by:

$$\nabla \epsilon(\mathbf{r}) \nabla \phi(\mathbf{r}) + 4\pi\rho(\mathbf{r}) = 0$$

$$\Delta G_{\text{electrostatic}} = {}^{1}/_{2} \sum_{i} q_i (\phi_i^{80} - \phi_i^1)$$

Theory and Methods: Nonpolar Energy Contributions

- Energy transferring small straight chain alkanes into water is linearly dependent on the alkane surface area
- Shown by this equation

$$\Delta G = \gamma A + b$$

- $A \rightarrow$ solvent accessible surface area
- γ, b $\rightarrow$ constants

Theory and Methods: Nonpolar Energy Contributions

- Experimental values of vacuum to water energies for alkanes used to obtain values γ, b
- Computation of free energies from equilibrium concentrations of solute in two phases given by

$$\Delta G = -RT \ln(C_1/C_2)$$

- $C_1/C_2 \rightarrow$ ratio expressed in units of molarity

Theory and Methods: Nonpolar Energy Contribution

- Curvature of solute-solvent is believed to affect the solute
- Given by following equation

$$\gamma(R)/\gamma(\infty) = 1/(1 \pm a/R)$$

- a → radius solvent, R → radius of solute-solvent surface
- Differences between curvatures of solvent and alkanes taken into account

Theory and Methods: Small Molecule Solvation Energy Calculations

- FDPB/γ method
- Resolution of 4.0 grids / Å
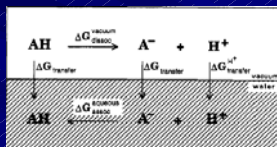
- Thermodynamic cycle used to determine transfer energies

$$\Delta G_{transfer}^{ion} = \Delta G_{transfer}^{neutral} \pm \Delta G_{transfer}^{H^+} \pm (\Delta G_{dissoc}^{vacuum} + \Delta G_{assoc}^{aqueous})$$

Theory and Methods: Calculated vs. Experimental data

- Comparison of amino acid N-methylacetamide
- ε = 1
- a) → OPLS parameters
- b) → *ab-initio* derived charges

Theory and Methods: PARSE

- Solvation energies for small molecules calculated reproduced
- Parameterize group constituents
- Total free energy is sum of constituent groups
- Nonpolar contributions calculated independently, subtracted later
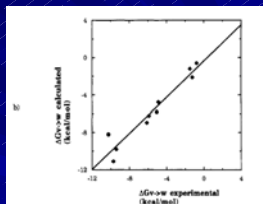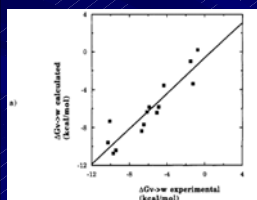- Gives 'experimental' electrostatic energy

Theory and Methods: PARSE

- Calculated electrostatic contribution best matched 'experimental' energies of functional group
- Fewest moments needed for alcohols, since O, H molecules known
- Complicated functional groups, i.e→ amide = ketone + amine
- Some not matched, i.e→ -COOH group

TABLE 1: Solvation Free Energies Using OPLS and Ab Initio-Derived Parameters[a]

		ΔG_{el}				$\Delta G_{vdw,np}$		
aacid	molecule	OPLS	ab initio	SA	$\Delta G_{c/v}$	OPLS	ab initio	$\Delta G_{solv,tot}^{b}$
arg	N-propylguanidine			307	2.46			-10.92
asn	acetamide	-12.73	-13.09	217	1.99	-10.73	-11.10	-9.72
asp	acetic acid	-10.35		213	1.97	-8.38		-6.70
cys	methylthiol	-5.26	-3.99	194	1.87	-3.39	-2.12	-1.24
gln	propionamide	-12.58	-11.94	249	2.16	-10.42	-9.79	-9.42
glu	propionic acid	-9.82		246	2.14	-7.68		-6.47
his	methylimidazole	-11.80	-10.40	260	2.22	-9.58	-8.19	-10.25
lys	N-butylamine	-5.85		274	2.29	-3.56		-4.38
met	methyl ethyl sulfide	-3.23	-3.42	261	2.22	-1.01	-1.20	-1.49
phe	toluene	-2.20	-3.00	294	2.39	0.19	-0.60	-0.76
ser	methanol	-8.16	-7.57	171	1.76	-6.41	-5.81	-5.08
thr	ethanol	-7.77	-6.72	208	1.95	-5.82	-4.77	-4.90
trp	methylindole	-8.44	-8.80	334	2.60	-5.84	-6.21	-5.91
tyr	p-cresole	-8.79	-9.39	305	2.45	-6.34	-6.95	-6.13
bb	N-methylacetamide	-9.52		255	2.19	-7.33		-10.08

- Calculations for polar amino acid side chain and backbone molecules
- $\varepsilon = 1$
- $\Delta G_{el} \to$ Electrostatic Free Energy
- SA→ Calculated Surface Area
- $\Delta G_{c/v} \to$ Calculated van der Waals free Energy
- $\Delta G_{tot} \to$ Total Free Energy

Results: Preoptimization

- Using OPLS and *ab-initio* parameters with $\varepsilon = 1$, results good
- Small errors
- 1.09 and 2.75 kcal/mole for OPLS
- 0.71 and 2.06 kcal/mole for *ab-initio*
- Results for $\varepsilon = 2$, less favorable

Results : Optimization

- Optimization for $\varepsilon=1$, same magnitude of error for OPLS, *ab initio*
- Greatly improved error for $\varepsilon=2$

TABLE 2: Solvation Free Energies Using Modified OPLS and *Ab Initio*-Derived Parameters[a]

		$\Delta G_{calc,\,tot}$				
aacid	molecule	OPLS inc crg[b]	OPLS red rad[c]	ab initio inc crg[d]	ab initio red rad[e]	$\Delta G_{exp,\,tot}$[f]
arg	N-propyl-guanidine					−10.92
asn	acetamide	−11.19	−9.89	−9.86	−11.11	−9.72
asp	acetic acid	−8.41	−7.17			−6.70
cys	methylthiol	−3.33	−1.57	−2.90	−1.47	−1.24
gln	propionamide	−10.85	−9.72	−9.01	−9.91	−9.42
glu	propionic acid	−7.84	−6.75			−6.47
his	methyl-imidazole	−9.81	−10.56	−7.69	−8.96	−10.25
lys	N-butylamine	−3.37	−3.56			−4.38
met	methyl ethyl sulfide	−1.11	−0.70	−2.48	−1.72	−1.49
phe	toluene	0.32	−0.37	−2.06	−1.62	−0.76
ser	methanol	−6.19	−5.62	−5.35	−5.04	−5.08
thr	ethanol	−5.52	−5.12	−4.73	−4.49	−4.90
trp	methylindole	−5.39	−5.45	−6.15	−6.45	−5.91
tyr	p-cresole	−6.02	−5.91	−6.60	−7.09	−6.13
bb	N-methyl-acetamide	−7.53	−7.01			−10.08

Results: Optimization of PARSE Parameter Set

TABLE 3: PARSE Group Charges, Solvation Free Energies, and Gas-Phase Dipole Moments[a]

A. Neutral Molecules

chemical group	atom	crg[b]	aacid	molecule[c]	ΔG_{el}	SA	$\Delta G_{el'}$	ΔG_{sol} calc	ΔG_{sol} exp[d]	ΔG_{corr}	calc μ	exp μ[e]	err μ
-OH	O	−0.49	ser	methanol	−7.21	158	1.77	−5.44	−5.08	0.36	1.74	1.71	0.03
	H	0.49	thr	ethanol	−6.97	190	1.95	−5.02	−4.90	0.12	1.81	1.69	0.12
				propanol	−6.95	220	2.11	−4.84	−4.83	0.01	1.81	1.68	0.13
				butanol	−6.87	249	2.27	−4.60	−4.72	0.12	1.82	1.66	0.16
				isopropyl alcohol	−6.69	216	2.09	−4.60	−4.76	0.16	1.74	1.66	0.08
				2-butanol	−6.51	246	2.25	−4.26	−4.58	0.32	1.73		
				3-methyl-1-butanol	−6.95	272	2.39	−4.56	−4.42	0.14	1.83		
NH₃	N	−0.90		ammonia	−5.68	119	1.56	−4.12	−4.31	0.19	1.33	1.47	0.14
	H₍₊₎	0.30											
-NH₂	N	−0.78		methylamine	−6.77	165	1.81	−4.96	−4.57	0.39	1.83	1.31	0.52
	H₁,₂	0.39		ethylamine	−6.58	194	1.97	−4.61	−4.50	0.11	1.77	1.22	0.55
				propylamine	−6.61	223	2.13	−4.48	−4.39	0.09	1.76	1.17	0.59
			lys	N-butylamine	−6.60	255	2.30	−4.30	−4.38	0.08	1.75	1.44[f]	0.31
-NH-	N	−0.50		dimethylamine	−6.53	201	2.01	−4.52	−4.29	0.23	1.84	1.03	0.81
	H	0.50		diethylamine	−5.97	194	2.35	−3.62	−4.07	0.45	1.83	0.92	0.91
-SH	S	−0.29	cys	methylthiol	−3.25	180	1.90	−1.35	−1.24	0.11	1.39	1.53	0.14
	H	0.29		ethylthiol	−3.21	210	2.06	−1.15	−1.30	0.15	1.42	1.58	0.16

- Calculated and experimental solvation energies
- Functional groups
- Excellent results for following groups, ketones, acids, amides

Results: Optimization of PARSE Parameter Set

B. Ionized Molecules

chemical group	atom	crg	aacid	molecule	ΔG_{el}	SA	ΔG_{cav}	calc	exp	ΔG_{err}	pKa	ΔG_s
NH4+	N	-0.32		ammonium	-70.71	145	1.71	-81.51	-81.53	0.02	9.25	196.4
	H, CH2(N)	0.33	lys+	N-butyl-ammonium	-71.80	258	2.36	-69.44	-69.24	0.20	10.60	210.6
-C-S-	C	-0.08	cys-	methylthiol ion	-78.54	177	1.88	-76.66	-76.79	0.13	10.30	350.6
	S	-0.92										
-COO-	C	0.10	asp-	acetate ion	-82.38	188	1.94	-80.44	-80.65	0.21	4.80	341.5
	$O_{1,2}$	-0.55	glu-	propionate ion	-81.32	219	2.11	-79.21	-79.12	0.09	4.88	340.3
aromatic ≥CO-	C_{ring}	-0.195	tyr-	p-cresole ion	-77.28	273	2.40	-74.88	-75.01	0.13	9.98	343.5
	H_{ring}	0.125										
	C(O)	-0.07										
	O	-0.50										
aromatic -NH-CH=	N	-0.35	his+	methyl-imidazolium	-66.57	239	2.21	-64.36	-64.13	0.23	6.95	216.6
N+H-CH=	H(N)	0.45										
	C	0.1413										
	H(C)	0.125										
-C-N3H-C-	$C_{1,2}$	0.35	arg+	N-p-guanidinium	-68.53	284	2.46	-66.07			13.65	
$(N_3H_2)_3^+$	N_c	-0.35										
	$H(N_a)$	0.45										
	N_a	-0.70										
	$H(N_a)$	0.40										

total ΔG_{err} 0.15   total μ err 0.36
aacid ΔG_{err} 0.10   aacid μ err 0.30

- Molecules used in PARSE
- Divided into chemical groups

---

Results: Optimization of PARSE Parameter Set

- Comparison of experimental and calculated free energies for amino acid side chains and NMA
- PARSE parameters used
- $\varepsilon = 1$
- Average and maximum errors for amino acid side chains and backbone is 0.10 and 0.23 kcal/mol

---
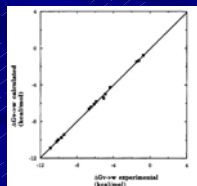
Results: Test of the PARSE parameters

- Comparison of calculated and experimental dipole moments for amino acid/backbone analogs
- Average absolute error for all molecules is 0.36D, amino acid/backbone was 0.30D
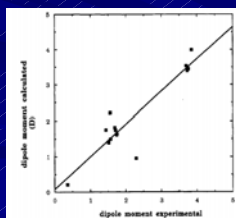
Discussion

- Able to calculate solvation free energies with solute polarizability
- Developed parameters to observe transfer energy of small molecules
- Includes backbones of amino acids
- Done by using FDPB method and PARSE

Discussion

- Future experiments involve testing methods and parameters on proteins
- Intra/inter molecular interactions taken into consideration
- Accuracy of small molecules allows for further study of proteins
- Protein-Protein or Protein-ligand interactions

Thank You!!!