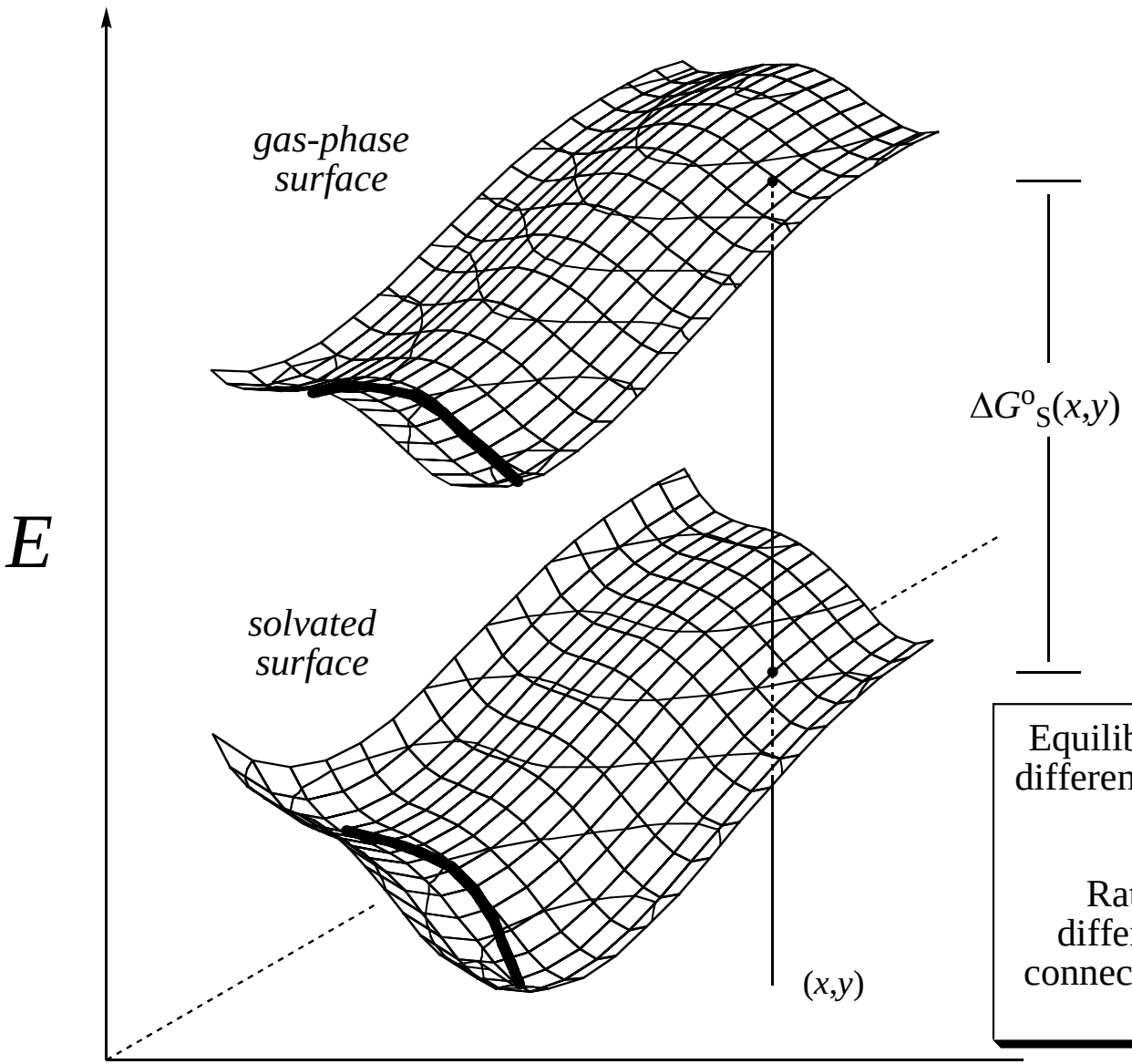
Modeling Environmental Fate Constants, Vapor Pressure, Partitioning, Solubility, and Environmentally Important Reaction Mechanisms

US EPA Computational Toxicology Forum
May 23, 2007

What Condensed Phases are Important to Environmental Chemistry?

- Homogeneous liquid solutions (the most common condensed phases found in experimental chemistry)
- Solids
- Surfaces
- Membranes

Related Potential Energy Surfaces



$\Delta G^{o}_{S}(x,y)$

Phase partitioning from separation between the surfaces themselves

Equilibrium constants from differences in energy of local minima

Rate constants from differences in energy of connected local minima and saddle points

Solvation Model 6 (SM6)

SM6 calculates aqueous solvation free energies based on gas or liquid-phase optimized geometries.

$$\Delta G_S^\circ = \Delta G_{ENP} + G_{CDS}$$

olarization
uclear
lectronic

change in the solute free energy due to electrostatic interactions between the solute and the bulk solvent and distortion of the solute's electronic structure in solution

The solvent is modeled as a dielectric continuum.

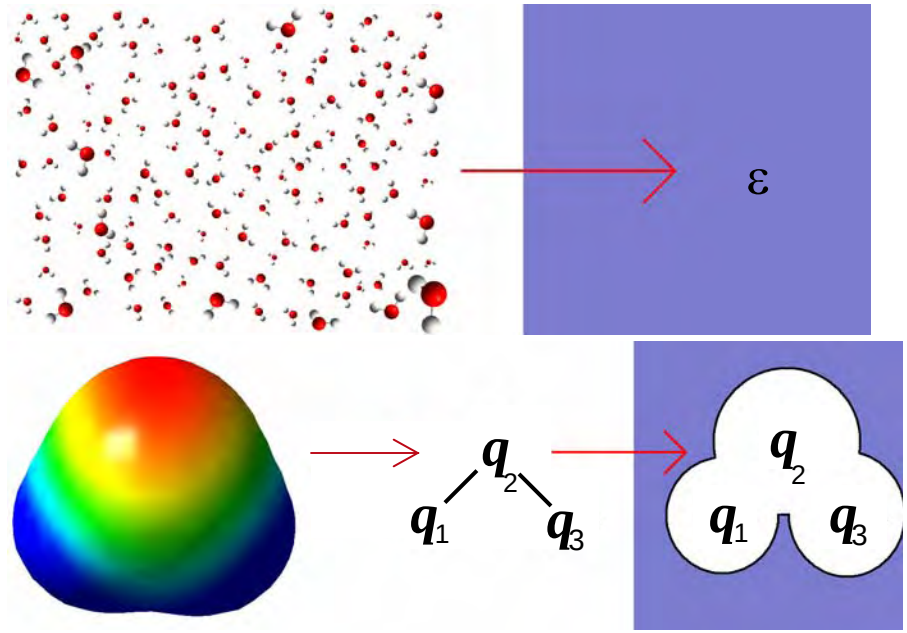
Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Theory Comput.* **2005**, *1*, 1133

Bulk Electrostatic Effects

Generalized Born (GB) equation

$$G_{\mathrm{P}} = -\frac{1}{2}\left(1-\frac{1}{\varepsilon}\right)\sum_{kk'}^{\text{atoms}} q_k \gamma_{kk'} q_{k'}$$

$$\gamma_{kk'} = \left(r_{kk'}^2 + \alpha_k \alpha_{k'} e^{-r_{kk'}^2 / d\alpha_k \alpha_{k'}}\right)^{-1/2}$$



Limiting behaviors…

$r_{kk'} >> 0$

$$G_{\mathrm{P}} = -\frac{1}{2}\left(1-\frac{1}{\varepsilon}\right)\frac{q_k q_{k'}}{r_{kk'}}$$

Coulomb's Law

$^1/_2$(–gas + solution)

$r_{kk'} = 0$

$$G_{\mathrm{P}} = -\frac{1}{2}\left(1-\frac{1}{\varepsilon}\right)\frac{q_k^2}{\alpha_k}$$

Born's Equation

monatomic ion

Solvation Model 6 (SM6)

SM6 calculates aqueous solvation free energies based on gas or liquid-phase optimized geometries.

$$\Delta G_S^\circ = \Delta G_{ENP} + G_{CDS}$$

tructure
ispersion
avitation

non-bulk electrostatic contributions to the free energy of hydration: first solvation shell effects

The G_{CDS} term is a parameterized term intended to minimize the deviation between the predictions and experiment.

It involves atomic surface tensions.

Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Theory Comput.* **2005**, *1*, 1133

Functional Form for G_{CDS}

geometry-independent part of the atomic surface tension of atom k

geometry of the solute

SM6

$$G_{\mathrm{CDS}} = \sum_{k}^{\text{atoms}} A_k \left(\sigma_k + \sum_{k'}^{\text{atoms}} \sigma_{kk'}(\mathbf{R}) \right)$$

atomic surface tension designed to account for the effect of solute geometry

solvent-exposed surface area of atom k

Microscopic Surface Tensions

$$G_{\mathrm{CDS}} = \sum_k A_k \sigma_k$$

Example: The SM*x* universal solvation models

$$\sigma_i = \hat{\sigma}_{Z_i}^{(n)} n + \hat{\sigma}_{Z_i}^{(\gamma)} \gamma + \hat{\sigma}_{Z_i}^{(\alpha)} \alpha + \hat{\sigma}_{Z_i}^{(\beta)} \beta$$

n is solvent index of refraction
γ is solvent *macroscopic* surface tension
α is solvent hydrogen-bonding acidity (Abraham)
β is solvent hydrogen-bonding basicity (Abraham)

Microscopic Surface Tensions

$$\Delta G^{\circ}_{\mathrm{S,expt}} - \Delta G_{\mathrm{ENP}} = G_{\mathrm{CDS}} = \sum_k A_k \sigma_k$$

Example: The SM*x* universal solvation models

$$\sigma_i = \hat{\sigma}^{(n)}_{Z_i} n + \hat{\sigma}^{(\gamma)}_{Z_i} \gamma + \hat{\sigma}^{(\alpha)}_{Z_i} \alpha + \hat{\sigma}^{(\beta)}_{Z_i} \beta$$

n is solvent index of refraction
γ is solvent *macroscopic* surface tension
α is solvent hydrogen-bonding acidity (Abraham)
β is solvent hydrogen-bonding basicity (Abraham)

After parameterization (about 74 parameters for 2400 data {H,C,N,O,F,S,P,Cl,Br-compounds} in 91 solvents including water) SM5.43 has a mean unsigned error of approximately 0.7 kcal mol^{-1} for neutrals and 5 kcal mol^{-1} for ions; this is roughly 5% for each case

SM6 Performance

Aqueous solvation free energy test set:

273 neutral solutes (containing H,C,N,O,F,P,S,Cl,Br)

112 ionic solutes (containing H,C,N,O,F,S,Cl,Br)

Average unsigned errors (kcal/mol)

	Ref.	neutrals	ions
DPCM/98	*Gaussian 98*	1.02	5.8
DPCM/03	*Gaussian 03*	1.40	13.0
IEF-PCM/03	*Gaussian 03*	1.10	8.6
CPCM/03	*Gaussian 03*	1.11	7.3
SM5.43	MN 2004	0.62	5.3

SM6 Performance

Aqueous solvation free energy test set:

273 neutral solutes (containing H,C,N,O,F,P,S,Cl,Br)

112 ionic solutes (containing H,C,N,O,F,S,Cl,Br)

Average unsigned errors (kcal/mol)

	Ref.	neutrals	ions
DPCM/98	*Gaussian 98*	1.02	5.8
DPCM/03	*Gaussian 03*	1.40	13.0
IEF-PCM/03	*Gaussian 03*	1.10	8.6
CPCM/03	*Gaussian 03*	1.11	7.3
SM5.43	MN 2004	0.62	5.3
SM6	**MN 2005**	**0.54**	**3.5**

SM5.43 for Organic Solvents

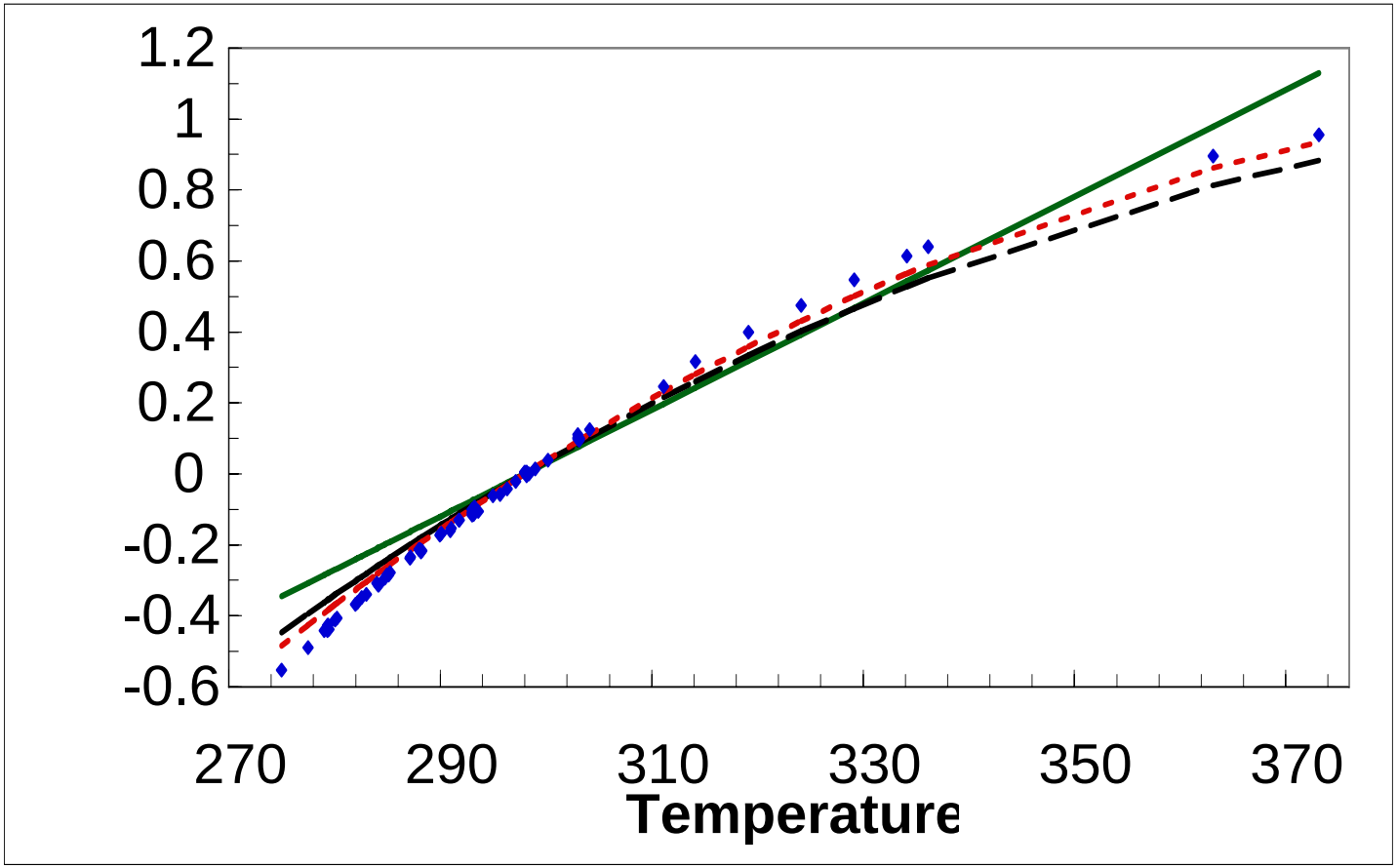
MUEs of the free energy of solvation in various solvents

		B3LYP/6-31G(d)	
Solvent	No. data	SM5.43R	C-PCM
acetonitrile	5	0.57	5.5
aniline	7	0.69	7.5
benzene	67	0.67	4.8
chlorobenzene	35	0.66	3.9
chloroform	94	0.55	4.6
cyclohexane	82	0.42	2.3
diethyl ether	60	0.76	2.4
dimethylsulfoxide	5	0.83	2.6
ethanol	5	0.94	2.9
heptane	59	0.40	2.7
nitromethane	5	0.75	3.2
tetrahydrofuran	5	0.48	3.5
toluene	46	0.47	3.5

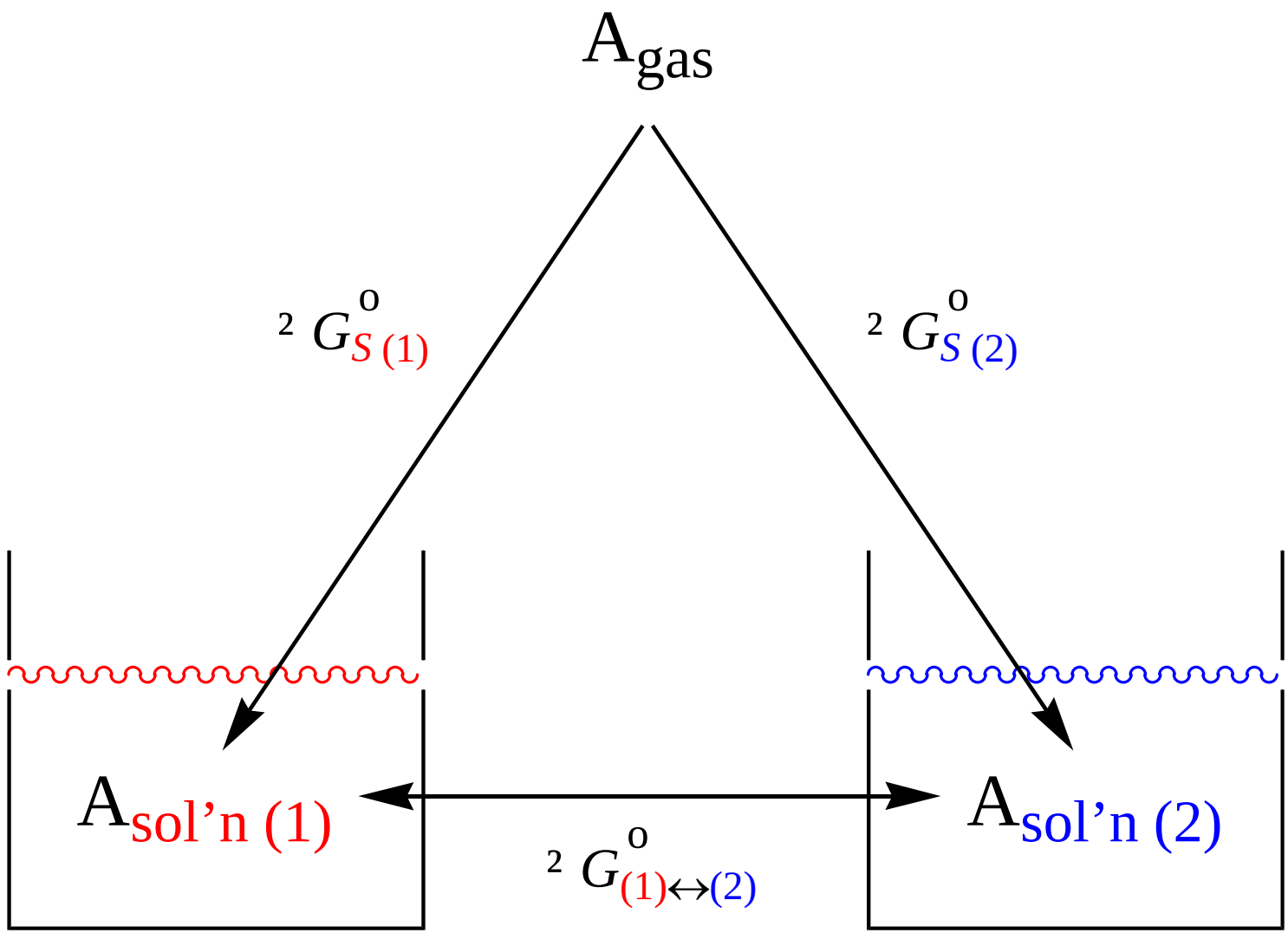
Temperature Dependence

Variation of the free energy of hydration of benzene relative to 298 K

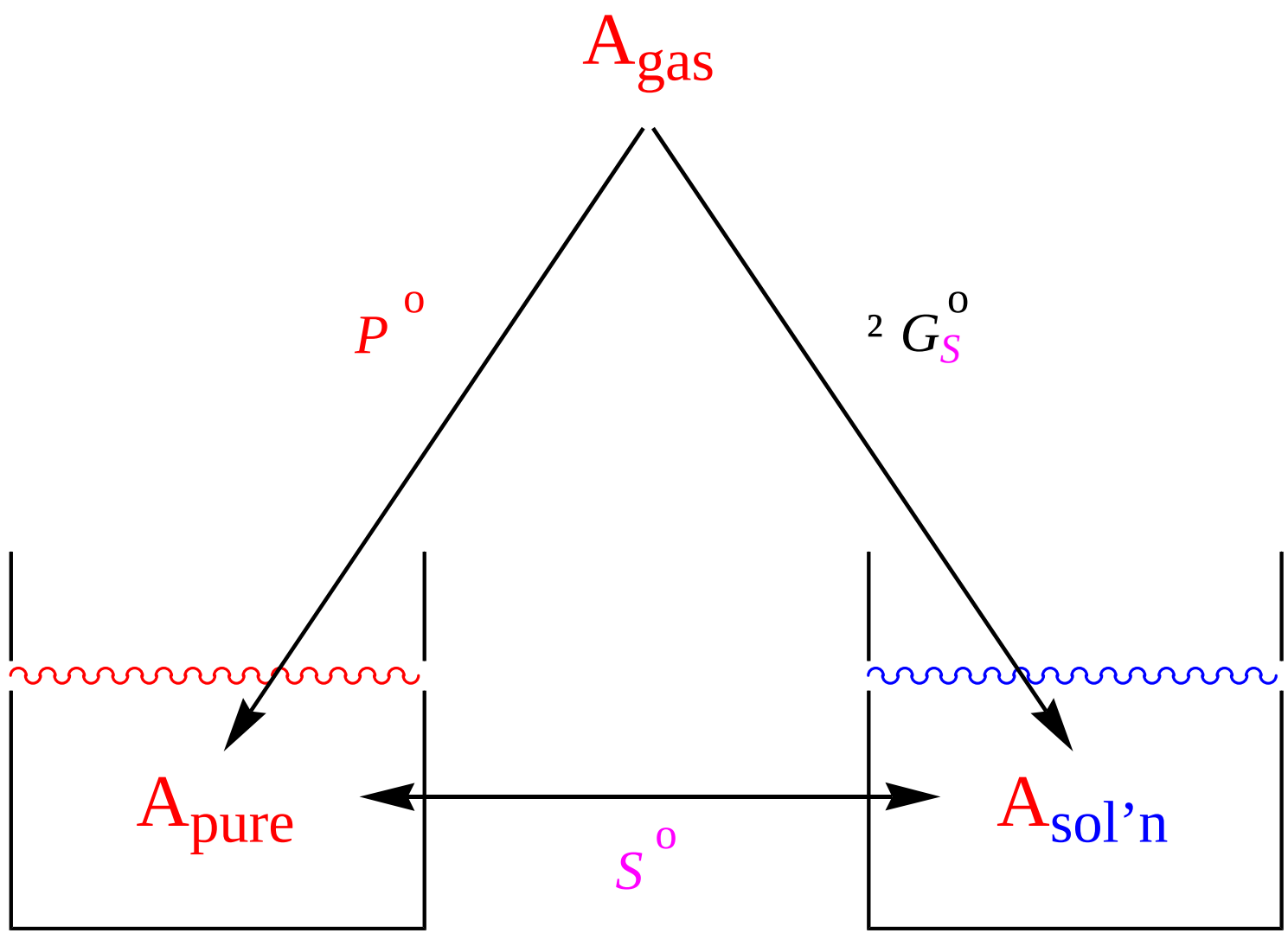
- Experiment
- Linear
- Unrestricted - 14 parameters
- Final - 7 parameters



Free Energies of Solvation and Partition Coefficients



Solubility from Solvation Free Energy and Vapor Pressure



Mean-unsigned error (MUE) in predicted log $P^{\bullet}$

		SM5.42R		
Solute class	No. data	HF	B3LYP	AM1
hydrocarbons	11	0.3	0.3	0.2
aromatics	6	0.2	0.3	0.2
alcohols/phenols	9	0.3	0.2	0.3
ethers	4	0.2	0.2	0.2
carbonyls	11	0.7	0.5	0.8
esters	7	0.3	0.2	0.6
CHN compounds	7	0.2	0.2	0.4
nitro compounds	5	0.1	0.2	0.3
HCNO compounds	60	0.3	0.3	0.4
halocarbons	15	0.4	0.4	0.5
all liquid solutes	**75**	**0.3**	**0.3**	**0.4**

Winget et al. *J. Phys. Chem. B* **2000**, *104*, 4726

Mean-unsigned error (MUE) in predicted log *S*

		SM5.42R			
Solute class	No. data	HF	B3LYP	AM1	UNIFAC
hydrocarbons	11	0.5	0.4	0.4	1.4
aromatics	6	0.1	0.0	0.1	0.2
alcohols/phenols	9	0.3	0.2	0.3	0.6
ethers	4	0.5	0.5	0.5	0.5
carbonyls	11	0.4	0.4	0.5	0.3
esters	7	0.3	0.3	0.2	0.1
CHN compounds	7	0.7	0.5	0.6	0.5
nitro compounds	5	0.4	0.4	0.2	0.2
HCNO compounds	60	0.4	0.4	0.4	0.6
halocarbons	10	0.2	0.3	0.2	0.4
all liquid solutes	**70**	**0.4**	**0.3**	**0.3**	**0.5**
solid solutes	**13**	**0.3**	**0.4**	**0.5**	**0.8**

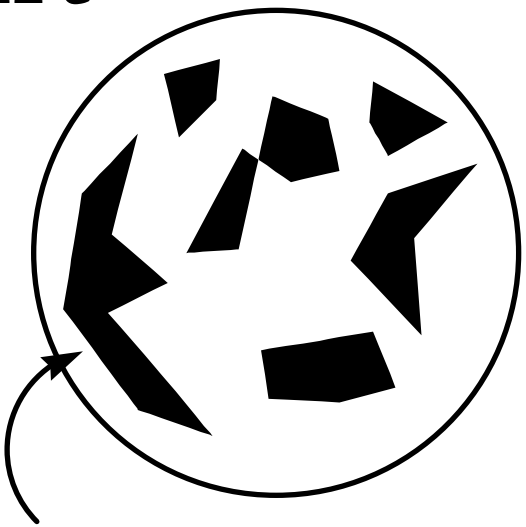
Thompson et al. *J. Chem. Phys.* **2003**, *119*, 1661

Soil/Water Partitioning

Important factor controlling the persistence of environmental contaminants

carbamates, phosphonothioates, polyhalogenated aromatics, ureas, horrible molecules

Dirt



organic carbon

$$K_{OC} = \frac{[X]_{soil} / \%OC}{[X]_{aqueous}}$$

remarkably constant from clay to loam to peat

“Solvent” Model for Soil

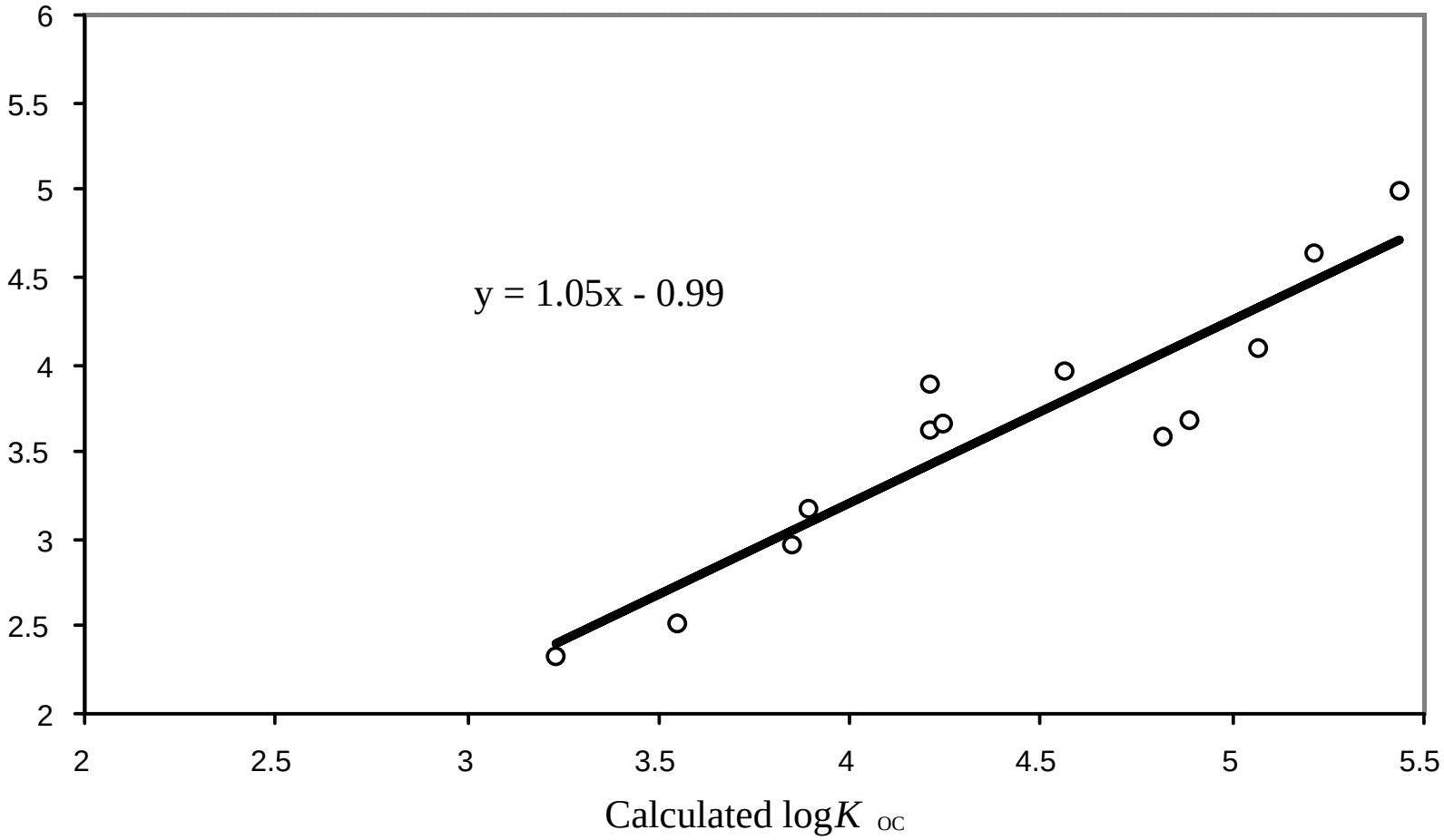
"Dirt/MIDI!"

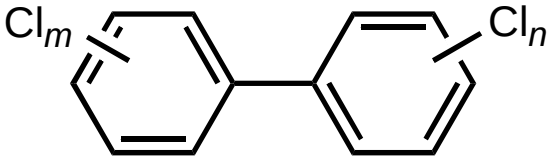
SM5.42R requires:	ε	n	α	β	γ
Best guess:	?	?	?	?	?
Regression fit:	(15.0)	1.379	0.61	0.60	46.0

Mean unsigned error over 387 compounds = 1 log unit

Winget et al. *Environ. Sci. Technol.* **2000**, *34*, 4733

Correlation of a Subset of Chlorinated Biphenyls

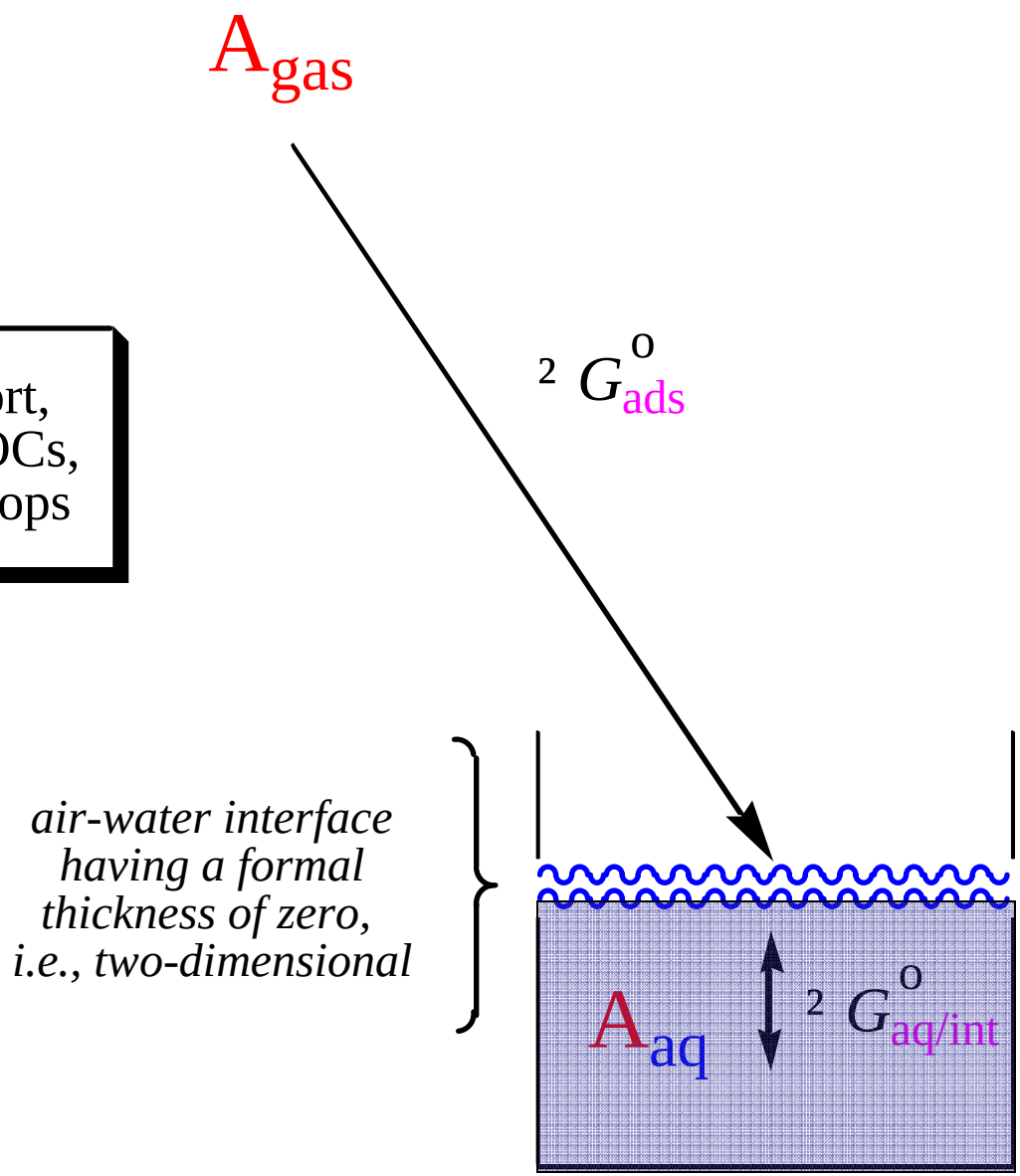




Theory allows for an atom-by-atom decomposition of the partitioning energies to better understand factors affecting them

Air/Surface Water Partitioning

Important in transport, e.g., stripping of VOCs, migration via fog drops



What About the Coupling Free Energy?

The SM5.0 Continuum Solvation Model

Coupling Term: **Proportional to Solvent-Accessible Surface Area**

Atomic surface tension functions σ fit to reproduce

^{2} G_S^o(expt) [over 2500 data]

SM5.0 **Each σ is a continuous differentiable function of:**

- **atomic number**
- **local geometry (connected atoms)**
- **<u>bulk solvent</u> properties including**
  - **index of refraction, surface tension**
  - **Abraham hydrogen bonding α and β**
  - **% "aromaticity/halogenicity"**

Surface Tension Functional Forms

$$\Delta G^{\mathrm{o}}_{\mathrm{coupling}} = \sum_i \sigma_i A_i$$

Universal functional form chosen to depend on *solvent* properties

n: index of refraction
γ: macroscopic liquid surface tension
α: Abraham hydrogen-bond donor parameter
β: Abraham hydrogen-bond acceptor parameter

$$\sigma_i = \hat{\sigma}^{(n)}_{Z_i} n + \hat{\sigma}^{(\gamma)}_{Z_i} \gamma + \hat{\sigma}^{(\alpha)}_{Z_i} \alpha + \hat{\sigma}^{(\beta)}_{Z_i} \beta$$

optimized parameters that define the SM5.0 model

SM5.0 for Bulk Water

Bulk descriptors are known from experiment:

$n = 1.342$
$\gamma = 71.2$
$\alpha = 0.82$
$\beta = 0.35$

Optimize σ parameters on 236 diverse molecule training set:

Mean unsigned error in free energies of solvation = 0.57 kcal mol^{-1}

Now, *freeze* the σ parameters

Can Surface Water be Characterized by Its Own "Solvent" Descriptors?

Use frozen σ parameters derived for bulk water:

Optimize bulk descriptors for 85 diverse molecule $K_{i/a}$ training set:

Surface	Bulk
n = 1.342	n = 1.342
γ = –144.6	γ = 71.2
α = 1.11	α = 0.82
β = 0.59	β = 0.35

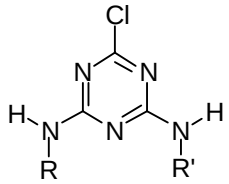
Mean unsigned error in $\log K_{i/a}$ = 0.47 log units

Possibly more important than error analysis is physical interpretation of the optimized parameters as they describe the water surface

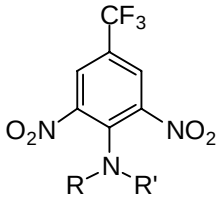
Error (log units) Analysis

Solute Class	SM5.0R-Surf Model	
1	7	0.72
2	4	0.40
3	3	0.29
4	3	0.40
5	2	-0.52
6	11	-0.32
7	5	0.22
8	7	-0.69
9	2	0.76
10	4	-0.16
11	5	-0.32
12	2	0.41
13	3	0.24
14	1	0.56
15	3	-0.76
16	15	-0.35
17	3	0.06
18	2	-0.06
19	3	0.13
	85	-0.07

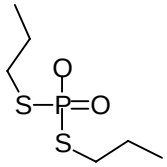
Model Can Be Used Independent of Experimental Data



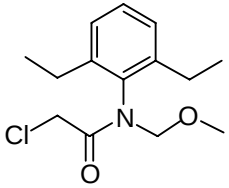
atrazine: R = iPr; R' = iPr
cyanazine: R = Et; R' = $C(CH_3)_2CN$
simazine: R = R' = Et



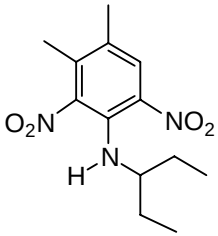
benfluralin: R = nBu; R' = Et
trifluralin: R = R' = nPr



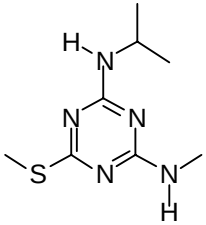
ethoprop



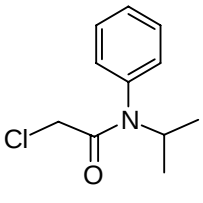
alachlor



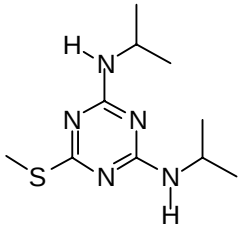
pendimethalin



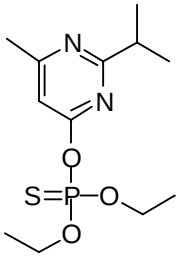
desmetryn



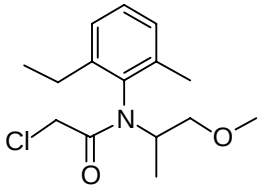
propachlor



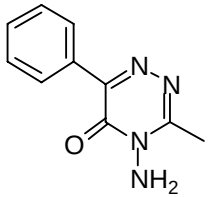
prometryn



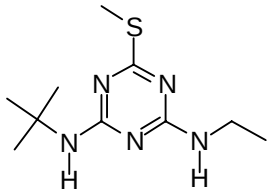
diazinon



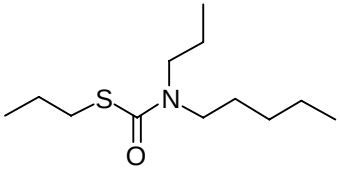
metolachlor



metamitron



terbutryn

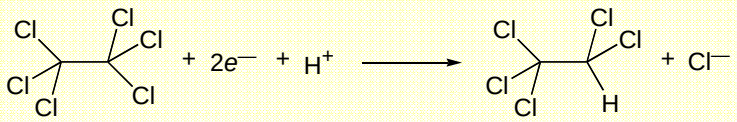
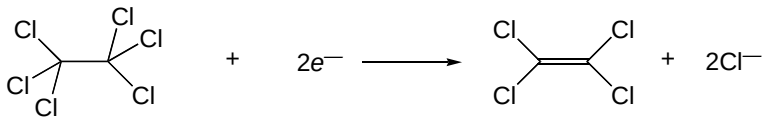
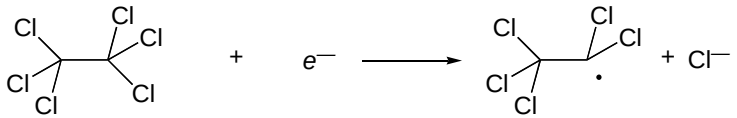


EPTC

Kelly et al. *J. Phys. Chem. A*, **2004**, *108*, 12882

Example: Reductive Dechlorination of Haloalkanes

Winget et al. *Theor. Chem. Acc.* **2004**, *112*, 217.

Reaction	E^o expt	E^o calc
$Cl^{\bullet} + e^- \longrightarrow Cl^-$	**2.54**	**2.45**
	**0.67**	**0.70**
	**1.15**	**1.00**
	**0.11**	**0.04**

$$\Delta G^{o'} = \Delta G^o + RT \ln \left[\frac{\dfrac{(1\,\mathrm{M\ C_2Cl_5H})(10^{-3}\,\mathrm{M\ Cl^-})}{(1\,\mathrm{M\ C_2Cl_6})(10^{-7}\,\mathrm{M\ H^+})(1\,\mathrm{M}\ e^-)^2}}{\dfrac{(1\,\mathrm{M\ C_2Cl_5H})(1\,\mathrm{M\ Cl^-})}{(1\,\mathrm{M\ C_2Cl_6})(1\,\mathrm{M\ H^+})(1\,\mathrm{M}\ e^-)^2}} \right]$$

$$= \Delta G^o + 5.5\ \mathrm{kcal\ mol^{-1}}$$

Protocol: E *from CCSD(T)/aug-cc-pVDZ//BPW91/aug-cc-pVDZ (thermal contributions from geometry level), solvation from SM5.42R/BPW91/DZVP, standard state p*H *7, all species 1 M but* Cl^- *buffered to 1 mM*

Example: Reductive Dechlorination of DDT

Lewis et al. *J. Chem. Ed.* **2004**, *81*, 596.

B3LYP/6-311+G*//B3LYP/6-31G*
+ SM5.42R/BPW91/6-31G*

$$\mathrm{DDT}_{(g)} + e^-_{(g)} \xrightarrow{\Delta G^o_{(g)}} \mathrm{DDX}_{(g)} + \mathrm{Cl}^-_{(g)}$$

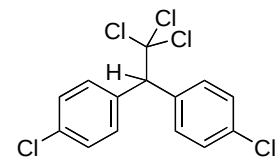
$\downarrow \Delta G^o_S(\mathrm{DDT})$ $\downarrow \Delta G^o = 0$ $\downarrow \Delta G^o_S(\mathrm{DDX})$ $\downarrow \Delta G^o_S(\mathrm{Cl}^-)$

$$\mathrm{DDT}_{(aq)} + e^-_{(g)} \xrightarrow[\mathbf{0.58\ V}]{\Delta G^o_{(1e)}} \mathrm{DDX}_{(aq)} + \mathrm{Cl}^-_{(aq)}$$

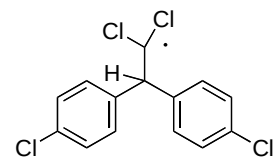
$$\mathrm{DDT}_{(g)} + \mathrm{H}^+_{(g)} + 2e^-_{(g)} \xrightarrow{\Delta G^o_{(g)}} \mathrm{DDD}_{(g)} + \mathrm{Cl}^-_{(g)}$$

$\downarrow \Delta G^o_S(\mathrm{DDT})$ $\downarrow \Delta G^o_S(\mathrm{H}^+)$ $\downarrow \Delta G^o = 0$ $\downarrow \Delta G^o_S(\mathrm{DDD})$ $\downarrow \Delta G^o_S(\mathrm{Cl}^-)$

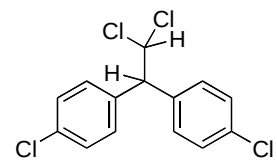
$$\mathrm{DDT}_{(aq)} + \mathrm{H}^+_{(aq)} + 2e^-_{(g)} \xrightarrow[\mathbf{0.86\ V}]{\Delta G^o_{(2e)}} \mathrm{DDD}_{(aq)} + \mathrm{Cl}^-_{(aq)}$$



DDT

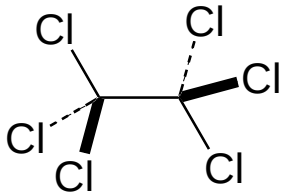


DDX



DDD

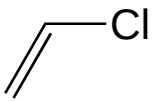
Example: Reductive Dechlorination of HCA



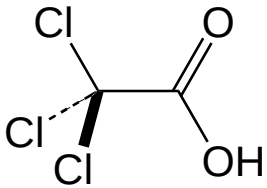
major industrial solvent (former dry-cleaning agent)

reducing power

reductive dechlorinations

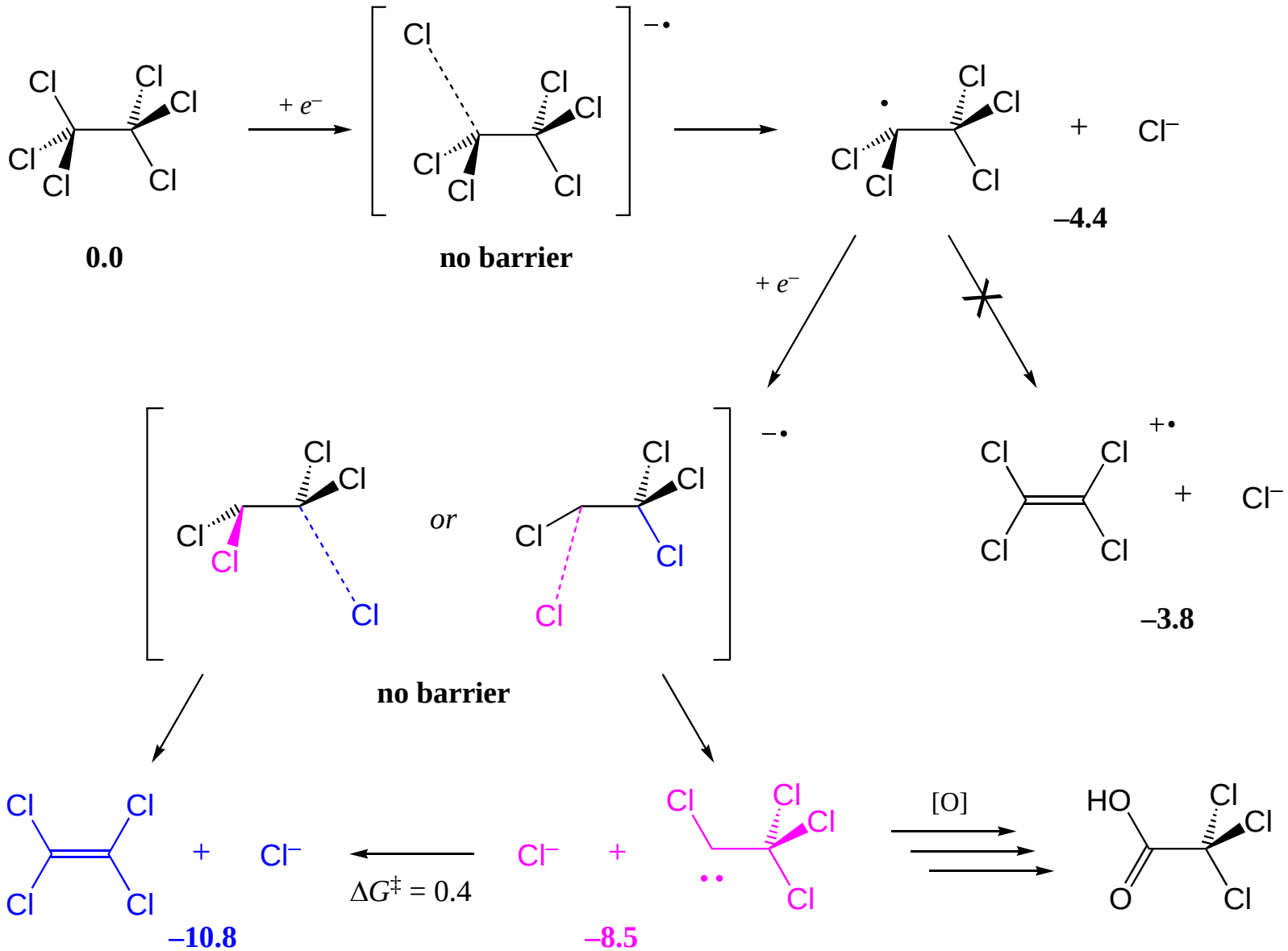


?

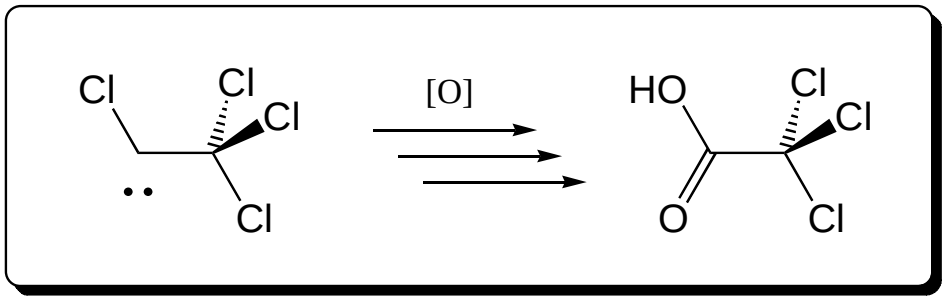


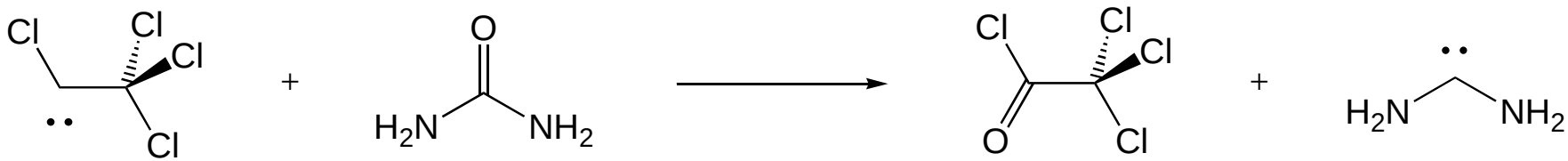
severe human health hazards

Example: Reductive Dechlorination of HCA

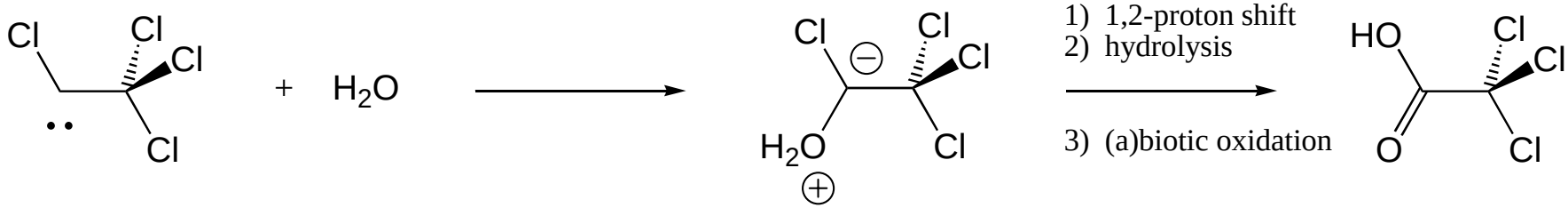


Example: Reductive Dechlorination of HCA





or



Acknowledgments

Biradicals and Diradicals

Dr. Fred Dulles
Dr. Sharon Worthington
Dr. Bradley Smith
Dr. Michael Sullivan
Dr. Mark Seierstad
Dr. William T. G. Johnson
Dr. Youngshang Pak
Dr. Stefan Debbert
Lee Solstad

Polynucleic Acid Dynamics

Dr. Maria Nagan
Dr. Ed Sherer
Stephanie Kerimo

Organometallics

Dr. Ben Gherman
Dr. Chris Kinsinger
Nathaniel Woodrum
Joe Scanlon
John Lewin
David Heppner

Solvation

Dr. Tianhai Zhu
Dr. Jiabo Li
Dr. James Xidos
Dr. Paul Winget
Casey Kelly
Adam Chamberlin

Senior Collaborators

Prof. Dan Falvey (Maryland)
Dr. Al French (USDA)
Prof. Daniel Liotard (Bordeaux)
Prof. Keiji Morokuma (Emory)
Prof. Eric Patterson (Truman State)
Prof. Larry Pratt (Fisk)
Prof. Bob Squires (Purdue, d.)
Prof. Bill Tolman (Minnesota)
Prof. Don Truhlar (Minnesota)
Dr. Eric Weber (US EPA, Athens)
Prof. Mike Zerner (Florida, d.)

Support from: ARO, NIST, NSF, EPA, Alfred P. Sloan Foundation, and Kodak

AMSOL, MN-GSM, HONDOPLUS, GAMESSPLUS, ZINDO-MN, etc. available from various sources

Introduction of Temperature Dependence into G_{CDS}

SM6

SM6T

$$G_{\mathrm{CDS}} = \sum_{k}^{\text{atoms}} A_k \left(\sigma_k + \sum_{k'}^{\text{atoms}} \sigma_{kk'}(\mathbf{R}) \right)$$

$$G_{\mathrm{CDS}} = G_{\mathrm{CDS}}(T_0) +$$

contribution at 298 K from SM6

$$\Delta S_S^{\circ}(T - T_0) +$$

$$\Delta C_P^{\circ}\left[(T - T_0) - T \ln(T/T_0)\right]$$

temperature dependence of the non-electrostatic contributions to the free energy of solvation relative to the value at T_0 (298 K)

Introduction of Temperature Dependence into G_{CDS}

SM6

SM6T

$$G_{CDS} = \sum_{k}^{atoms} A_k \left(\sigma_k + \sum_{k'}^{atoms} \sigma_{kk'}(\mathbf{R}) \right)$$

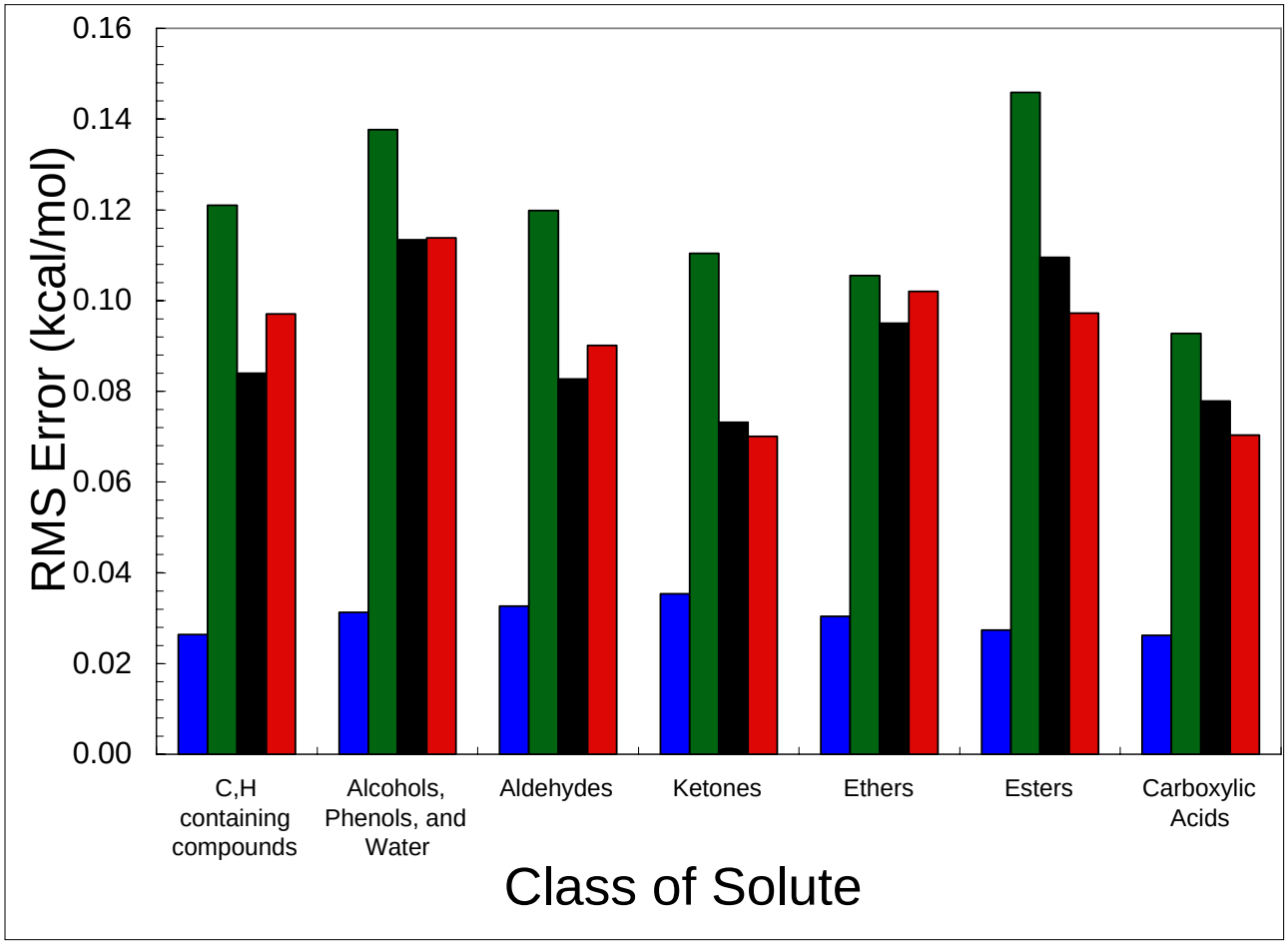
$$G_{CDS} = G_{CDS}(T_0) + \left(\sum_{k}^{atoms} A_k \left(\sigma_k^{\Delta S} + \sum_{k'}^{atoms} \sigma_{kk'}^{\Delta S}(\mathbf{R}) \right) \right)(T - T_0) + \left(\sum_{k}^{atoms} A_k \left(\sigma_k^{\Delta C_P} + \sum_{k'}^{atoms} \sigma_{kk'}^{\Delta C_P}(\mathbf{R}) \right) \right)\left[(T - T_0) - T \ln(T/T_0)\right]$$

entropy-like component

heat capacity-like component

temperature dependence of the non-electrostatic contributions to the free energy of solvation relative to the value at T_0 (298 K)

Errors



RMS Errors over entire database

(■) Individual fits ~ 0.03

(■) van't Hoff ~ 0.12

(■) Unrestricted ~ 0.09

(■) Final ~ 0.09

Null Hypothesis ~ 0.7

Chamberlin, A. C.; Cramer, C. J.; and Truhlar, D. G. *J. Phys. Chem. B*, **2006**, *110*, 5665

How to Compute Free Energy of Adsorption?

$$\Delta G^{o}_{\text{ads}} = \Delta G^{o}_{\text{3D}\rightarrow\text{2D}} + \Delta G^{o}_{\text{coupling}}$$

Most physically meaningful when concentration effects on translational free energy are eliminated through choice of standard state—coupling then measures only interaction free energy between solute and surface water

Accounts for loss of entropy associated with reduction in spatial dimensionality and depends on standard-state concentration choice for the different phases

Measured experimentally as $-RT \ln ([A]^{o}_{\text{int}} / [A]^{o}_{\text{air}})$

Statistical Mechanics for Dimensionality Free Energy Change 1

Molar translational partition function Z

$$Z = \frac{1}{N_A!}\left(\frac{L^\circ}{\Lambda}\right)^{MN_A}$$

where N_A is Avogadro's number, L^o is a standard-state length unit (e.g., meter), M is the number of dimensions, and Λ is the de Broglie wavelength

$$\Lambda = \left(\frac{2\pi mkT}{h^2}\right)^{-\frac{1}{2}}$$

where m is the solute molecular mass, k is Boltzmann's constant, T is the temperature, and h is Planck's constant

Statistical Mechanics for Dimensionality Free Energy Change 2

$$Z = \frac{1}{N_A!}\left(\frac{L^\circ}{\Lambda}\right)^{MN_A}$$

$$U^\circ = kT^2\left(\frac{\partial \ln Z}{\partial T}\right)_{N,V} \qquad H^\circ = U^\circ + PV \qquad S^\circ = k\ln Z + kT\left(\frac{\partial \ln Z}{\partial T}\right)_{N,V}$$

$$U^\circ = \frac{M}{2}RT \qquad PV = RT \qquad S^\circ = R\ln\left[\frac{e^{(M+2)/2}}{N}\left(\frac{L^\circ}{\Lambda}\right)^M\right]$$

For "average" molecular weight, dimension change costs 14 to 15 kcal mol^{-1}

$$\Delta G^\circ_{3D\to 2D} = \Delta H^\circ_{3D\to 2D} - T\Delta S^\circ_{3D\to 2D}$$

$$\Delta G^\circ_{3D\to 2D} = -RT\ln\left(\frac{\Lambda}{L^\circ}\right).$$

Application to air/water surface: Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. B* **2004**, *108*, 12882.