Why is it so hard to design ligands to bind to proteins?

George M. Whitesides

Harvard University
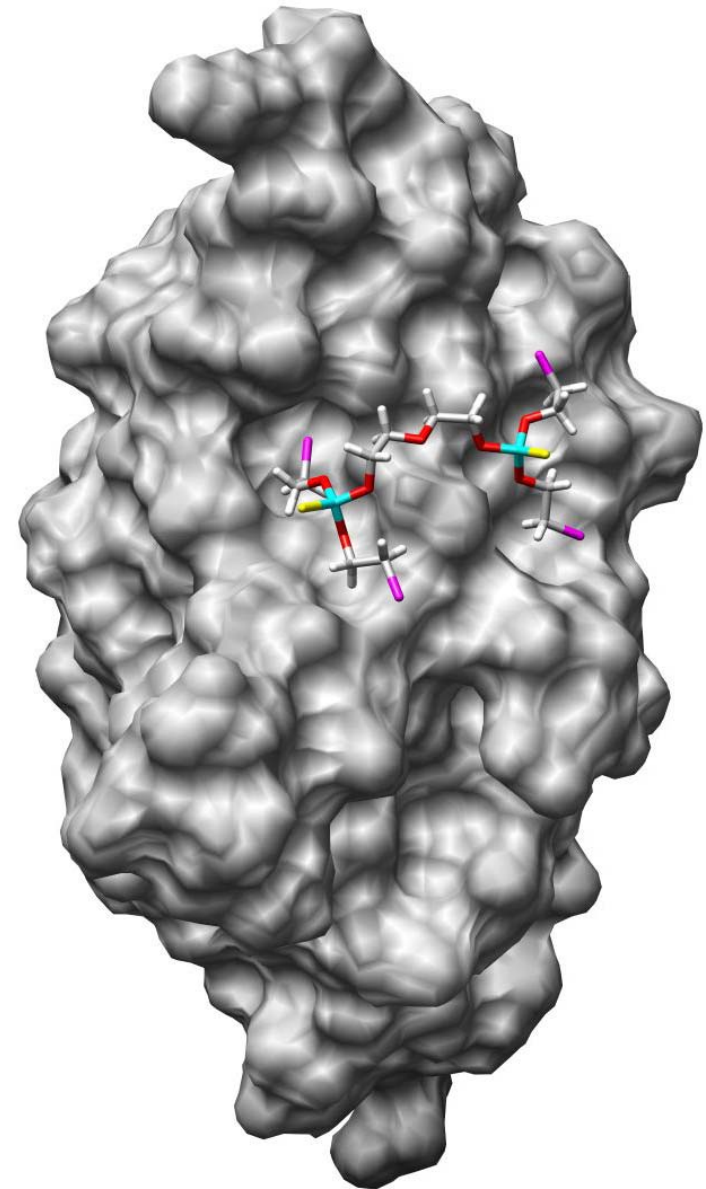
Department of Chemistry and Chemical Biology

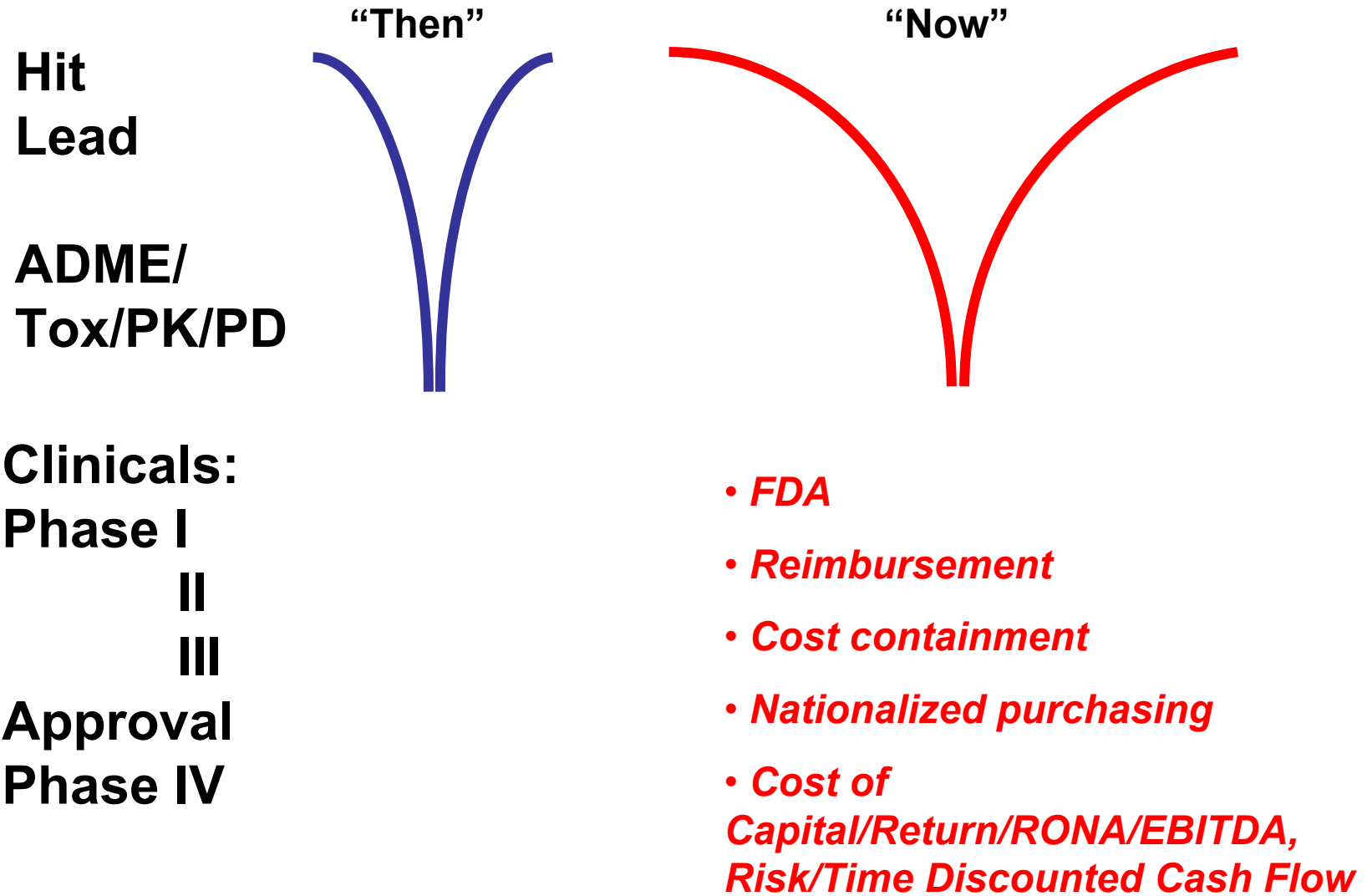gwhitesides@gmwgroup.harvard.edu

Molecular Recognition

- Non-covalent chemistry

- Ubiquitous

  - Protein-Ligand; Protein-Protein; Channels; Enzymatic Catalysis (ground/transition state) etc

- G, H, and S are all important

- Chemistry has good intuition about H; poor intuition about S

Hit / Lead to Drug

Hit
Lead

ADME/
Tox/PK/PD

Clinicals:
Phase I
      II
       III
Approval
Phase IV

"Then"

"Now"

- *FDA*

- *Reimbursement*

- *Cost containment*

- *Nationalized purchasing*

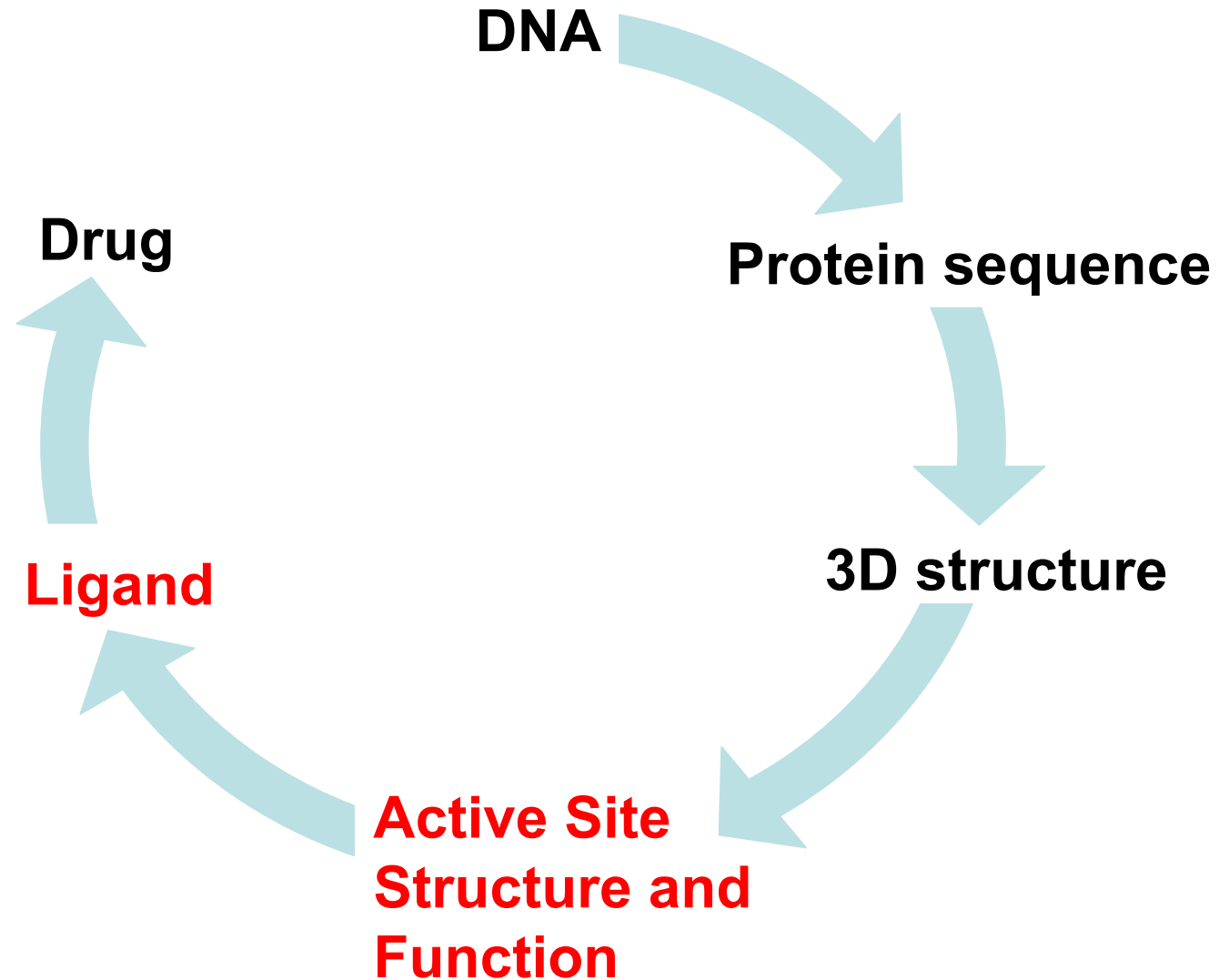- *Cost of Capital/Return/RONA/EBITDA, Risk/Time Discounted Cash Flow*
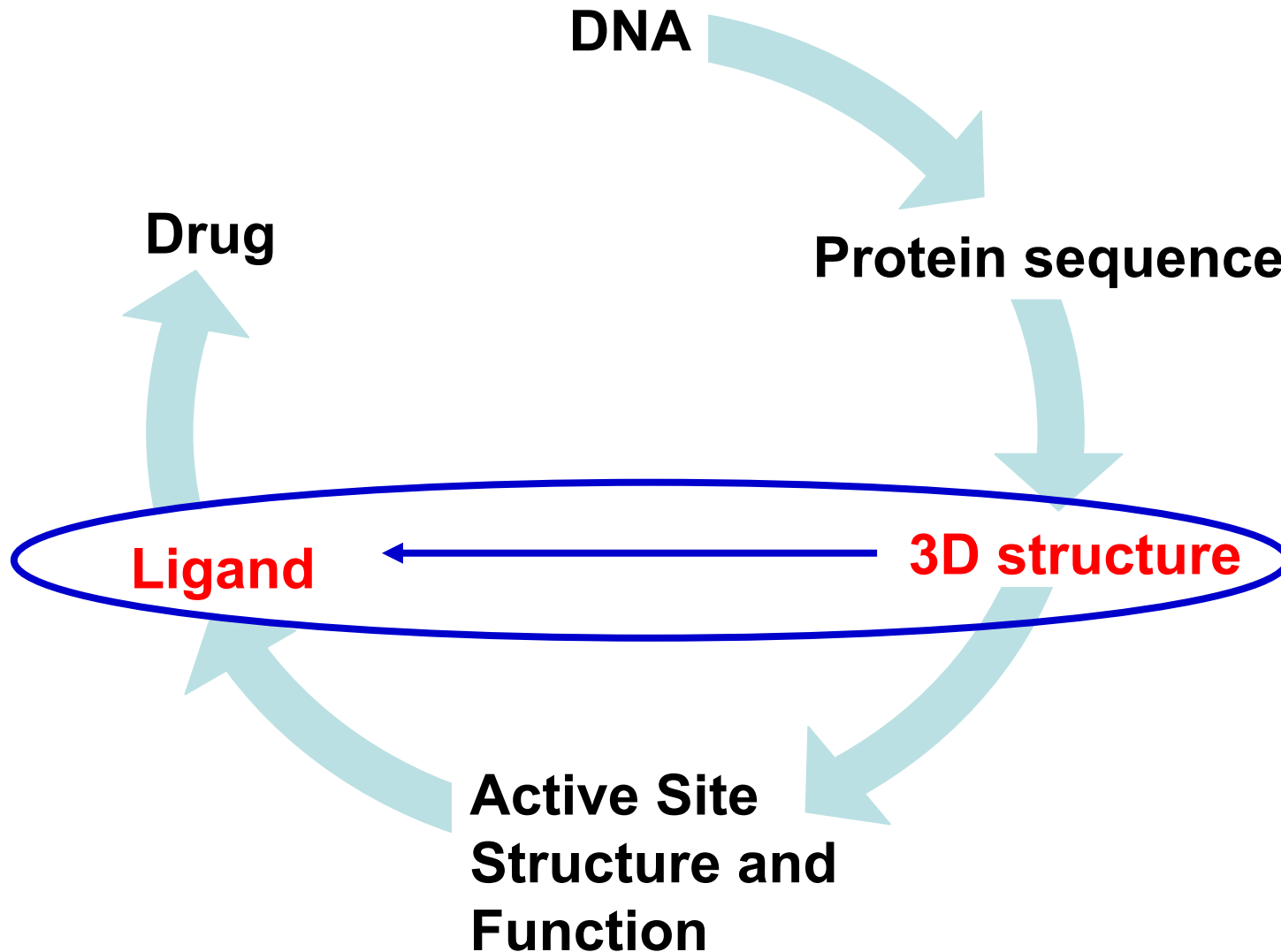
Why can't we design drugs?

- Drugs are much more complicated than ligands.

- Ligands are hard enough.

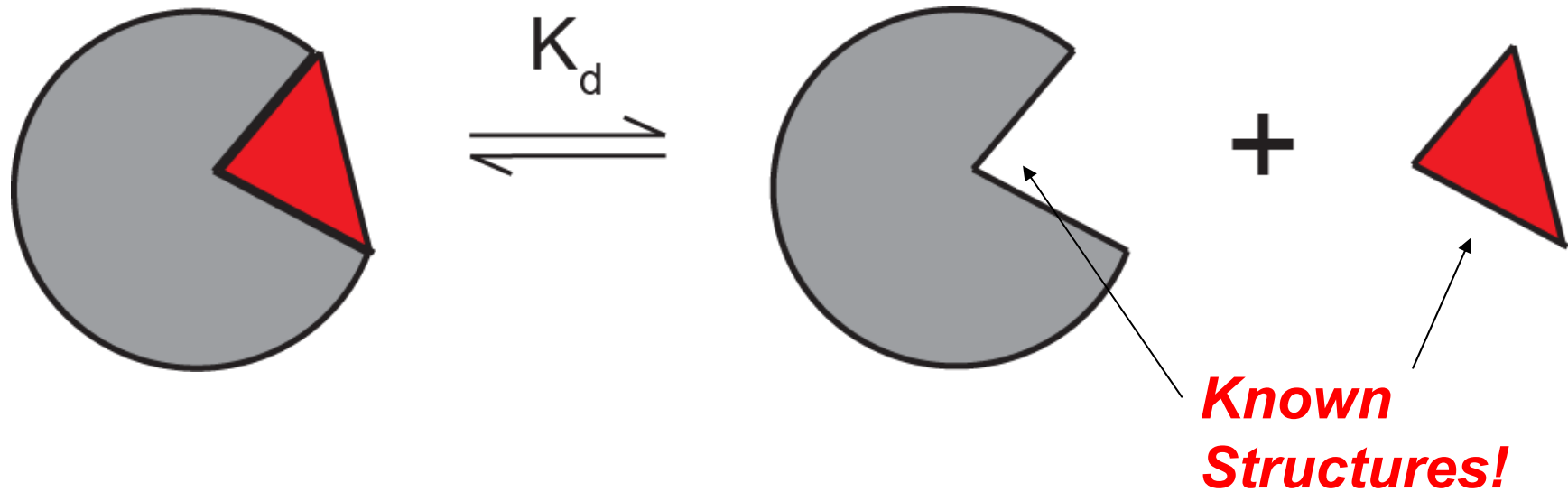- **Biology is all molecular recognition and networks.**
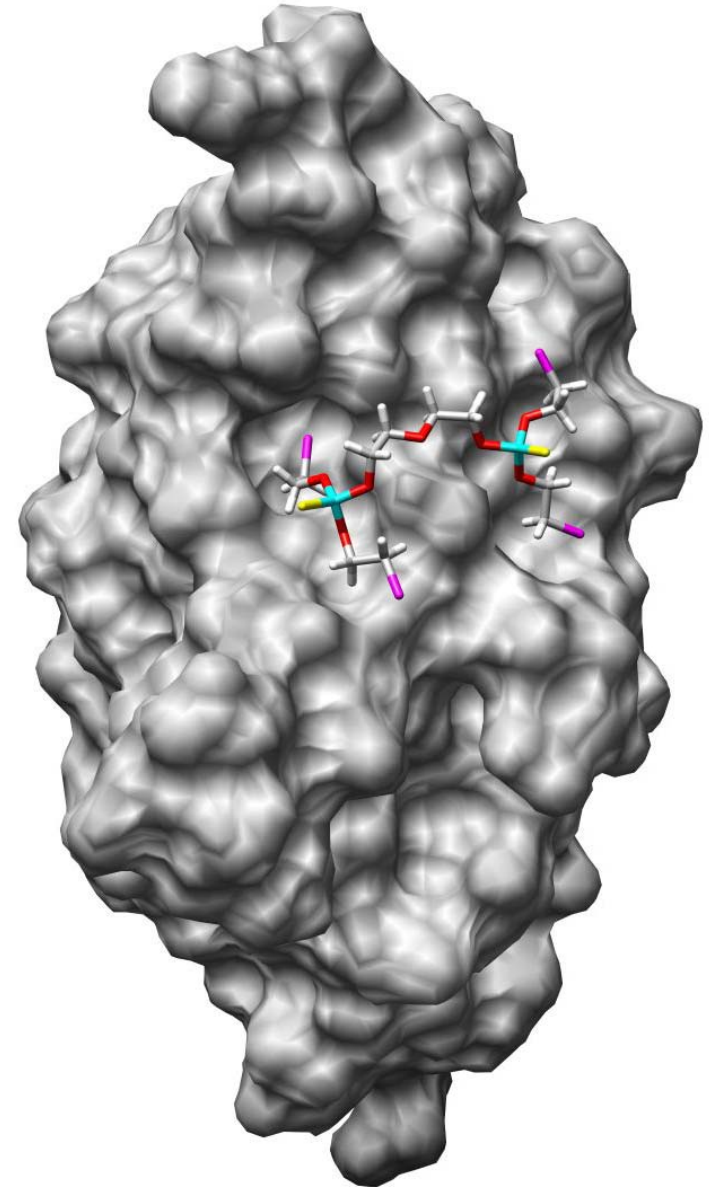
Rational Drug Design: Gene to Drug
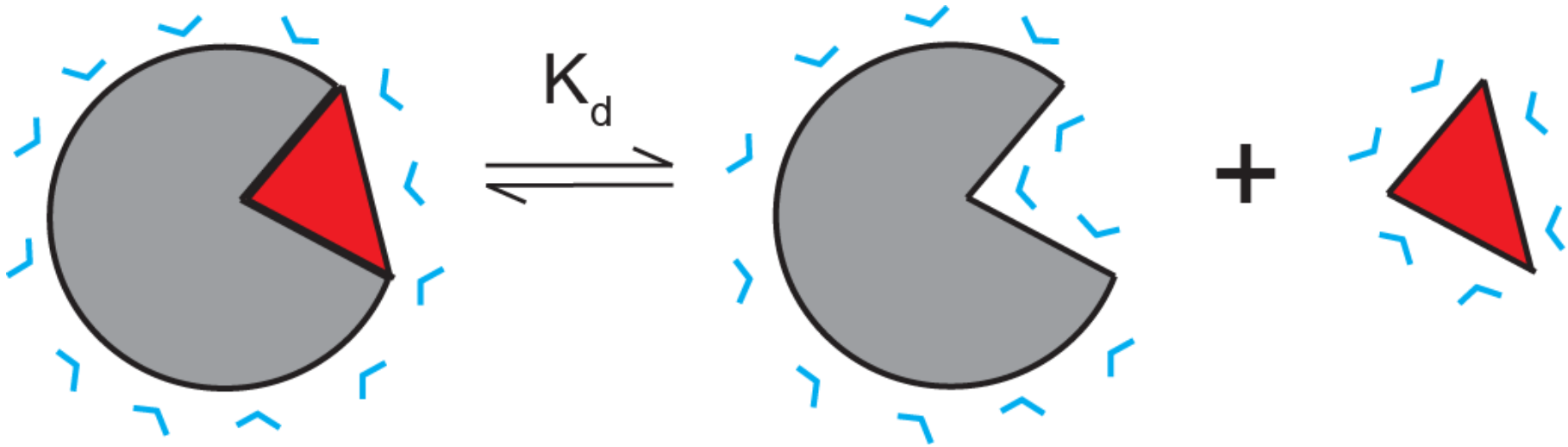
Rational *Ligand* Design: *Protein Structure* to *Lead*

Why is this so hard to predict?

$$K_d = \frac{[\text{⬤}][\text{▲}]}{[\text{⬤}]} \ (M)$$
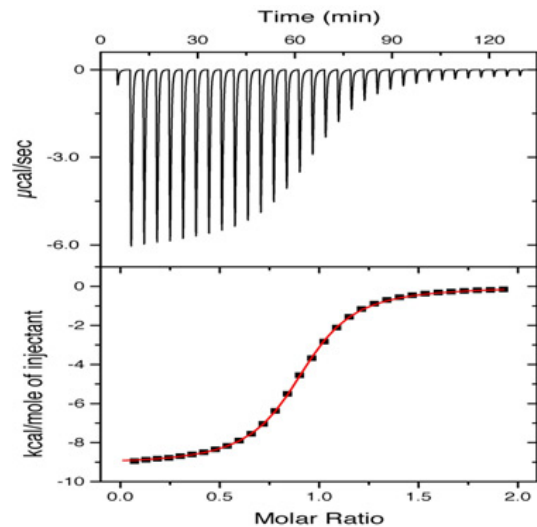
Known Structures!

Water!

$$K_d = \frac{[\,\text{◗}\,][\,\text{◀}\,]}{[\,\text{◕}\,]} \ (M)$$
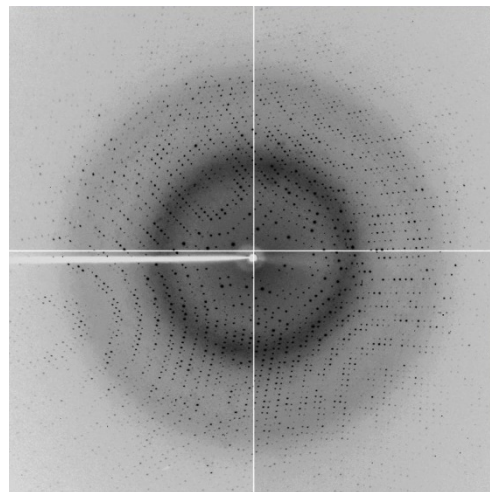
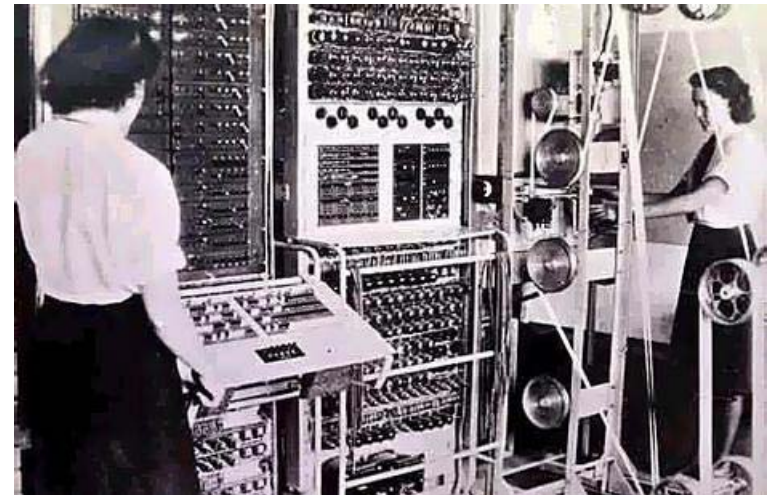$$K_d = e^{-\Delta G^0 / RT} = e^{-(\Delta H^0 - T\Delta S^0)/RT}$$

What Has Changed?

calorimetry   **x-ray crystallography**   **computational modeling**

ITC to Measure Enthalpy (and Infer Entropy)

- **Van't Hoff analysis**

$$\ln K_d = \frac{\Delta H^\circ}{R}\left(\frac{1}{T}\right) - \frac{\Delta S^\circ}{R}$$

  - ΔH° and ΔS° both vary with T
  - Protein structure is dependent on T
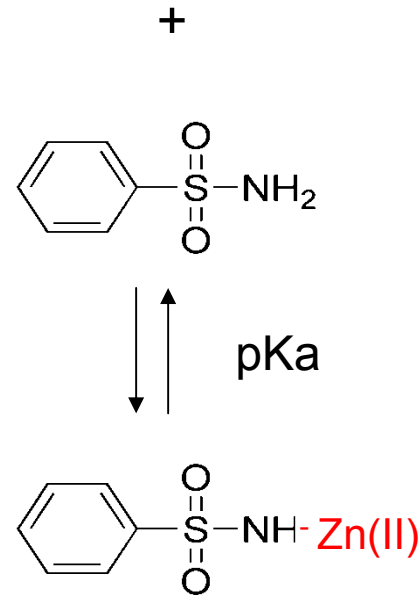 - Water structure depends on T

- **Isothermal titration calorimetry (ITC)**
  - Direct measurement of heat released upon binding
  - Constant temperature
  - Commercial instruments are available
  - Can estimate ΔC_p from variation of ΔH° with T

Representative Data from ITC
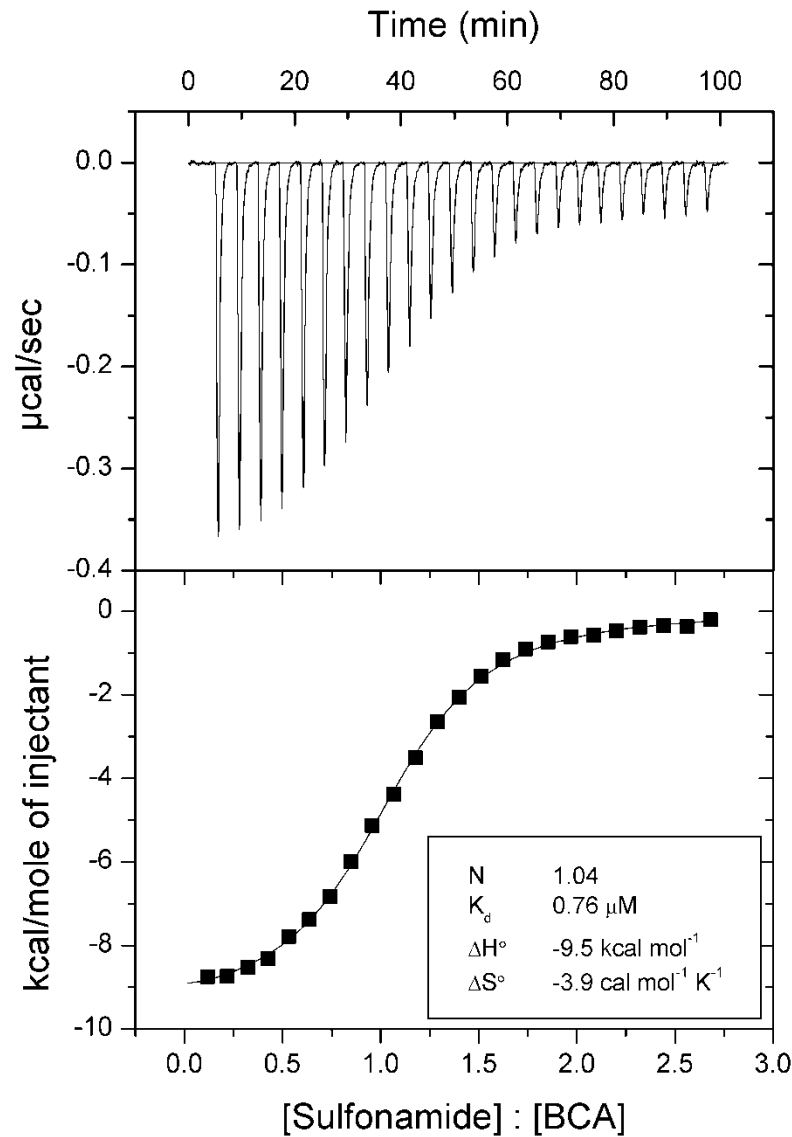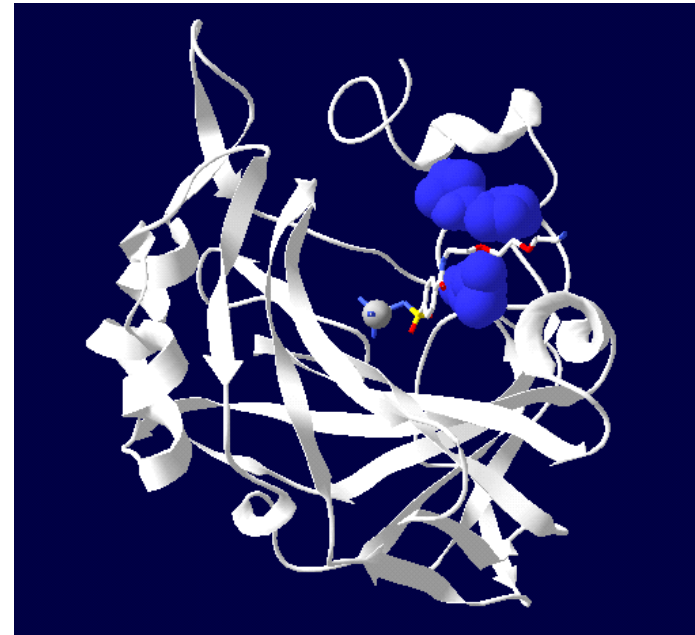

Bovine Carbonic Anhydrase II

Limitations: ITC

• *Requires significant protein*

• *Ligand must be soluble*

• *Obviously limited when q is small*

Molecular Recognition in Water: Protein-Ligand Interactions

- G = H – TS
- The Hydrophobic Effect
- Protein Plasticity
- "Induced Fit" (H and S)
- "Sloppy fit," rather than "lock-and-key"
- Other Interactions: Charge Networks
- Entropy/enthalpy (H/S) compensation
- Water, Solvation, and Entropy
- Association of OH$^-$ and non-polar surface

Carbonic Anhydrase: A Model Protein

- **Commonly used model protein for physical-organic studies**

- **Stable ($T_m = 65\ ^\circ C$)**

- **Monomeric, 30 kDa**

- **No disulfide bonds**

- **Structure is dominated by 10 β-sheets**

- **Zn(II)-OH cofactor in active site**

- **Function: CO_2 hydration**

- **Binding of sulfonamide inhibitors is well-characterized**
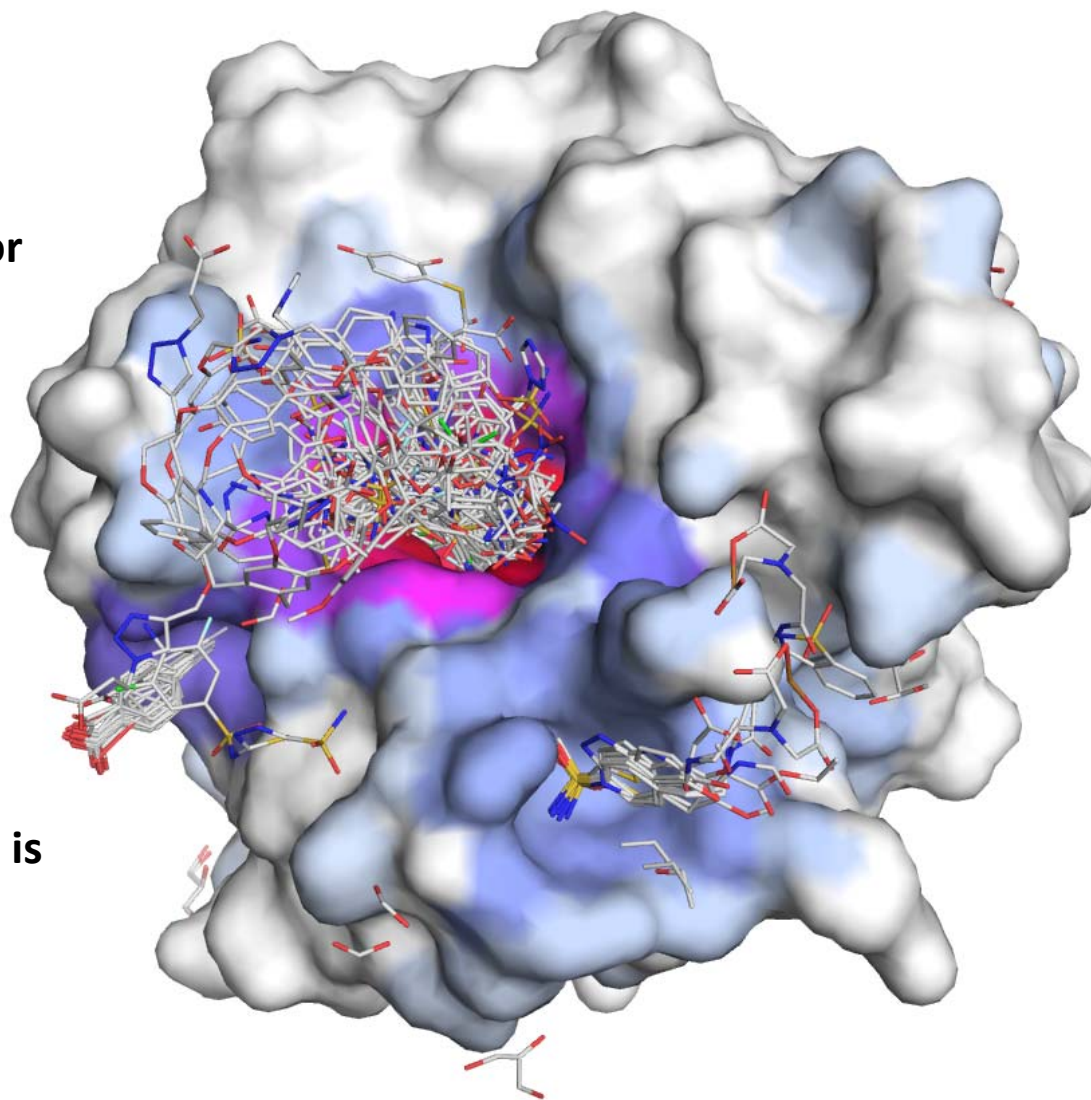
The Position of Aryl Sulfonamide Ligands is Restricted by Binding to the Zn(II)
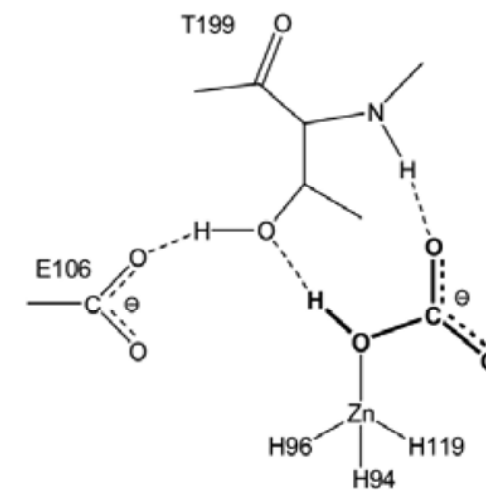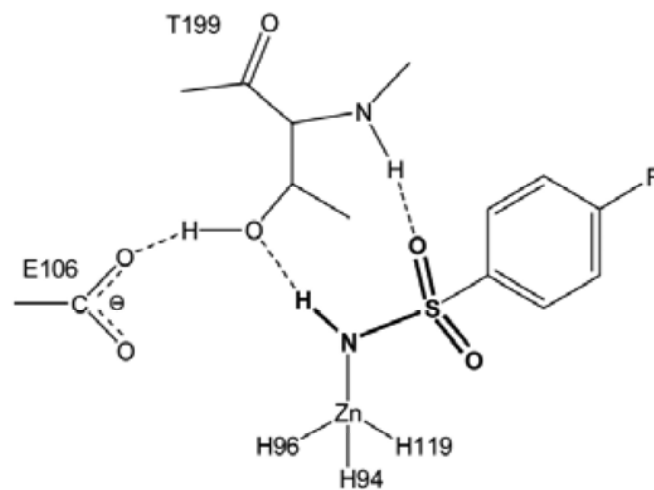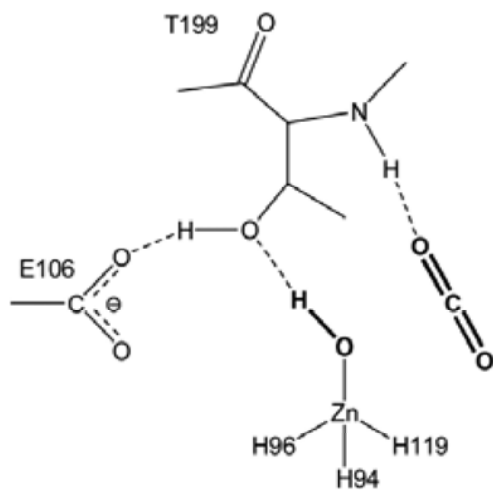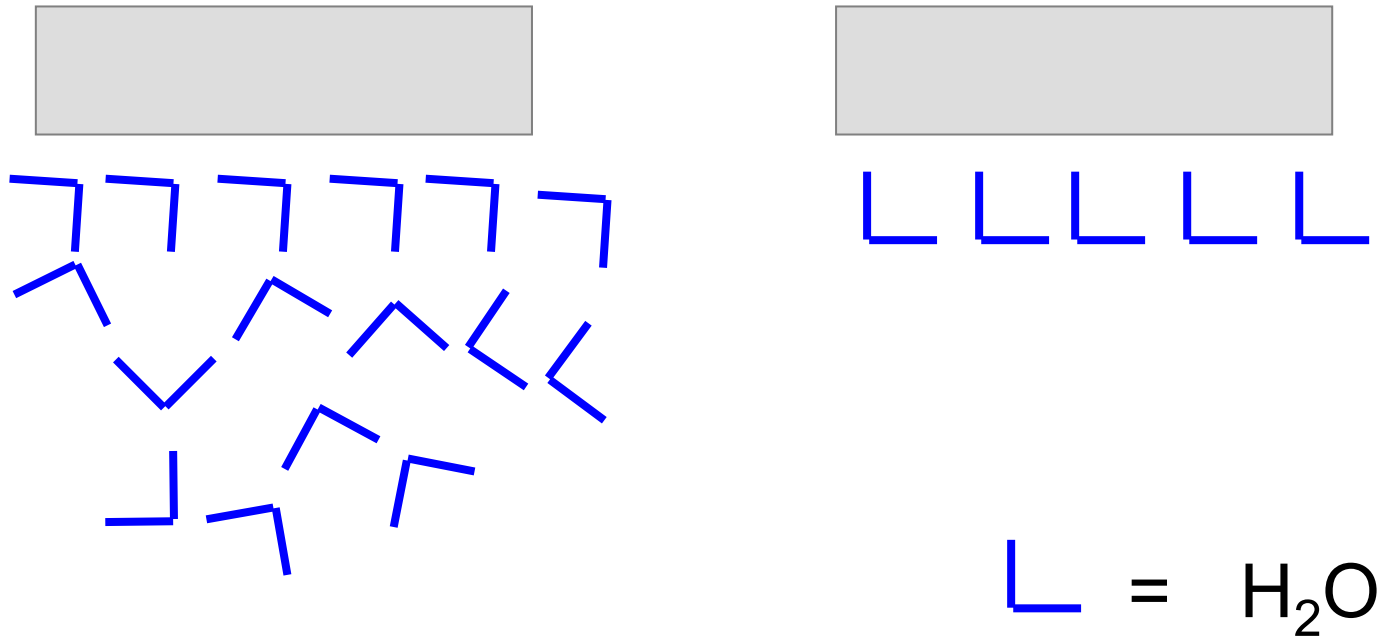
Water

- **High polarity**
- **Structured, but how?**
  - Hydrogen Bonding
- **How does structure translate to entropy?**
- **High, temperature-dependent, dielectric constant [ε = f(T)]**
- **Small partial molar volume ($\overline{V}_m$)**
- **High surface tension. Free energy required to form a cavity in water:**

$$\Delta G = \gamma \times \Delta A$$

where γ : **surface tension**

   ΔA : **change in surface area necessary to hold the solute**

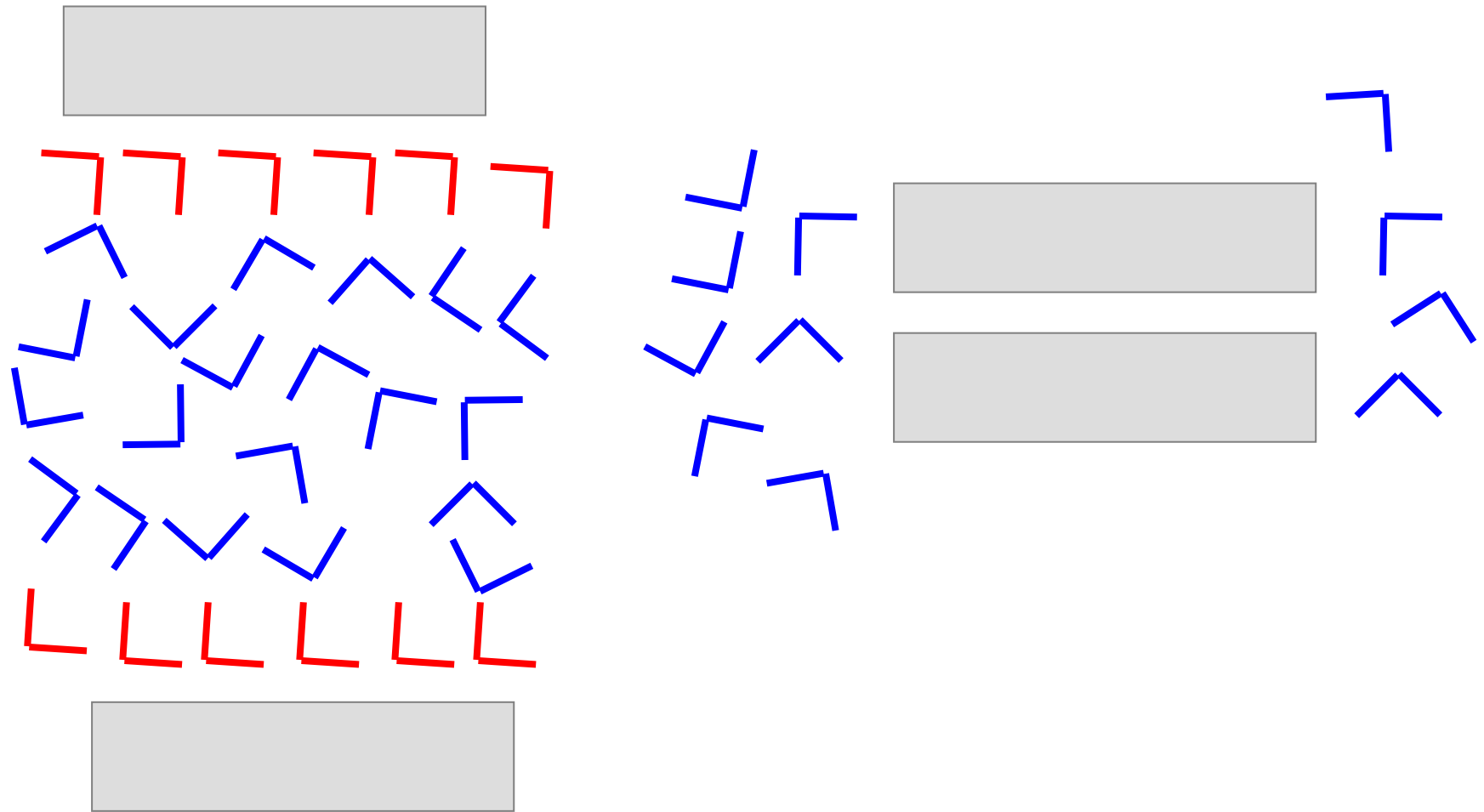# The Kauzmann-Tanford (KT) Hydrophobic Effect

∟ = H_2O

Orientation of near-surface water?

Structure of ordered layer?

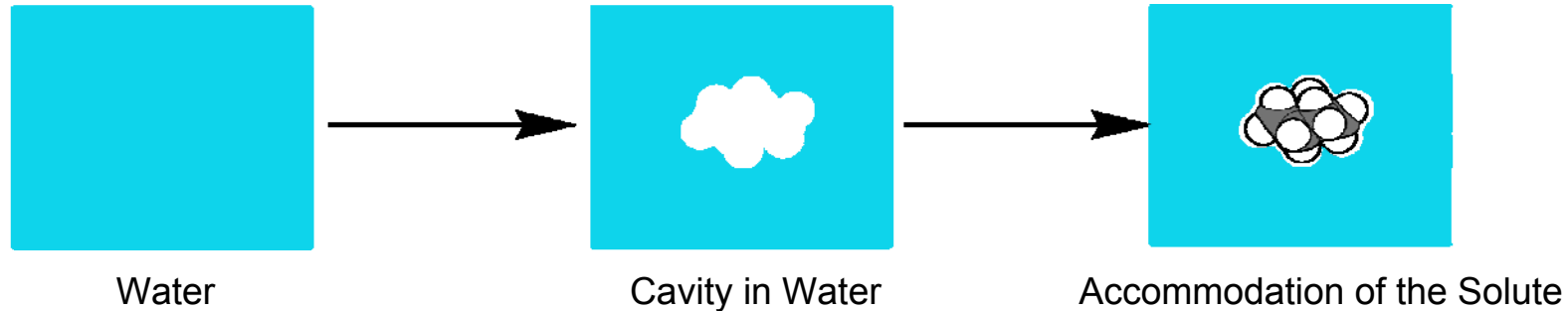**Structure of bulk water? (3 or 4 H-bonds?)**

KT Model: Contact of Hydrophobic Surfaces Releases Structured Water-- Dominated by Entropy
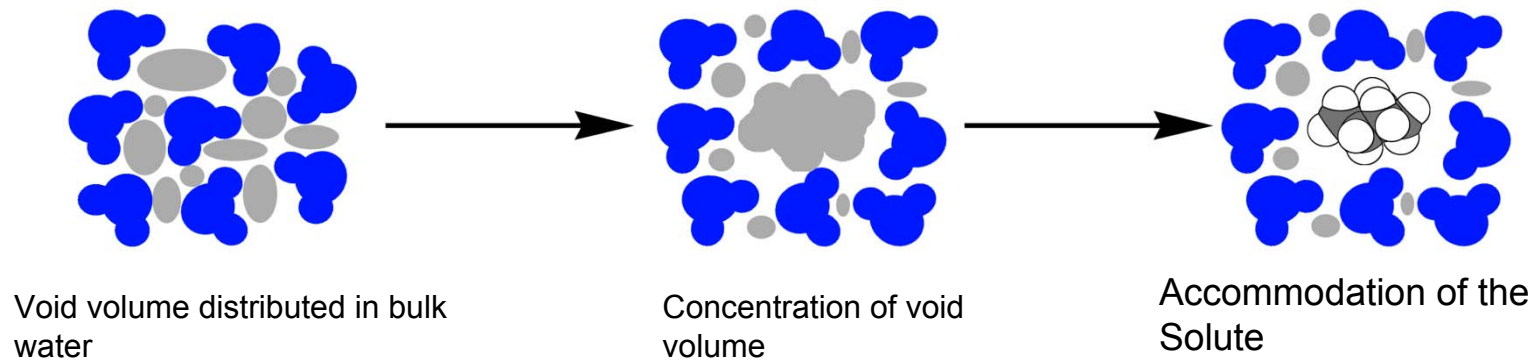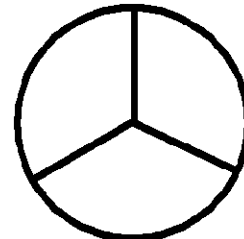
Hydrophobic Effect

- ## Cavity Formation Models

Surface Tension Model: $\Delta G = \gamma \times \Delta A$ where γ : surface tension, ΔA : change in surface area to hold the solute

| Water | Cavity in Water | Accommodation of the Solute |

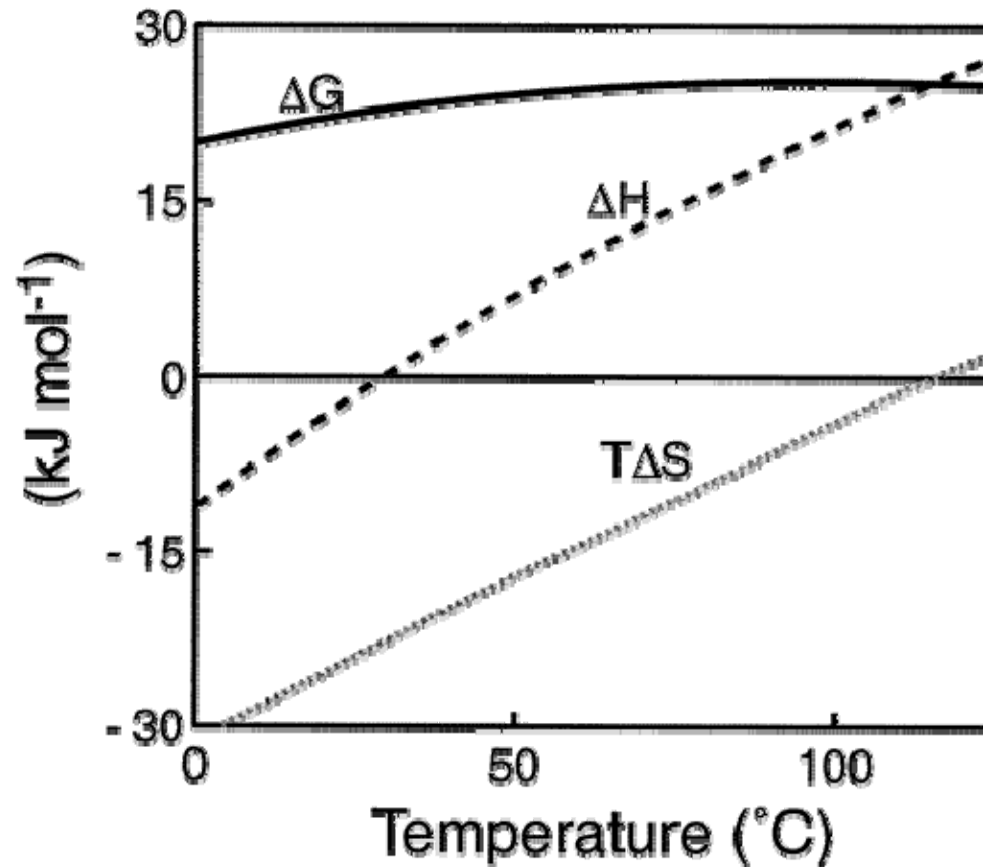Void Volume Aggregation Model: $\Delta G = T\Delta S$ of arranging many small void volume elements into a large volume

Void volume distributed in bulk water

Concentration of void volume

Accommodation of the Solute

- ## Mercedes Benz (MB) Model

K. Dill and coworkers *J. Phys. Chem. B* **2002**, *106*, 521-533

Topography of the Hydrophobic Surface

Hydrophobic Effect –
Enthalpy/Entropy Compensation

K. Dill and coworkers *J. Phys. Chem. B* **2002**, *106*, 521-533.

Molecular Recognition in Water:
Protein-Ligand Interactions

Molecular Recognition in Water: Protein-Ligand Interactions, or water?

Molecular Recognition in Water: Protein-Ligand Interactions

- G = H – TS
- The Hydrophobic Effect
- Protein Plasticity
- "Induced Fit" (H and S)
- "Sloppy fit," rather than "lock-and-key"
- Other Interactions: Charge Networks
- Entropy/enthalpy (H/S) compensation
- Water, Solvation, and Entropy
- Association of OH$^-$ and non-polar surface

Our Experimental Strategy

- **Use the simplest possible design**
- **Perturbation of the ligand (leave the protein = CA) constant**
- **Compare aliphatic and aromatic groups**
- **Compare "hydrophobicity" in binding to ligand, and in partitioning**
- **Test data for statistical significance**
- **Get X-ray structures for everything**
- **Compare with theory**
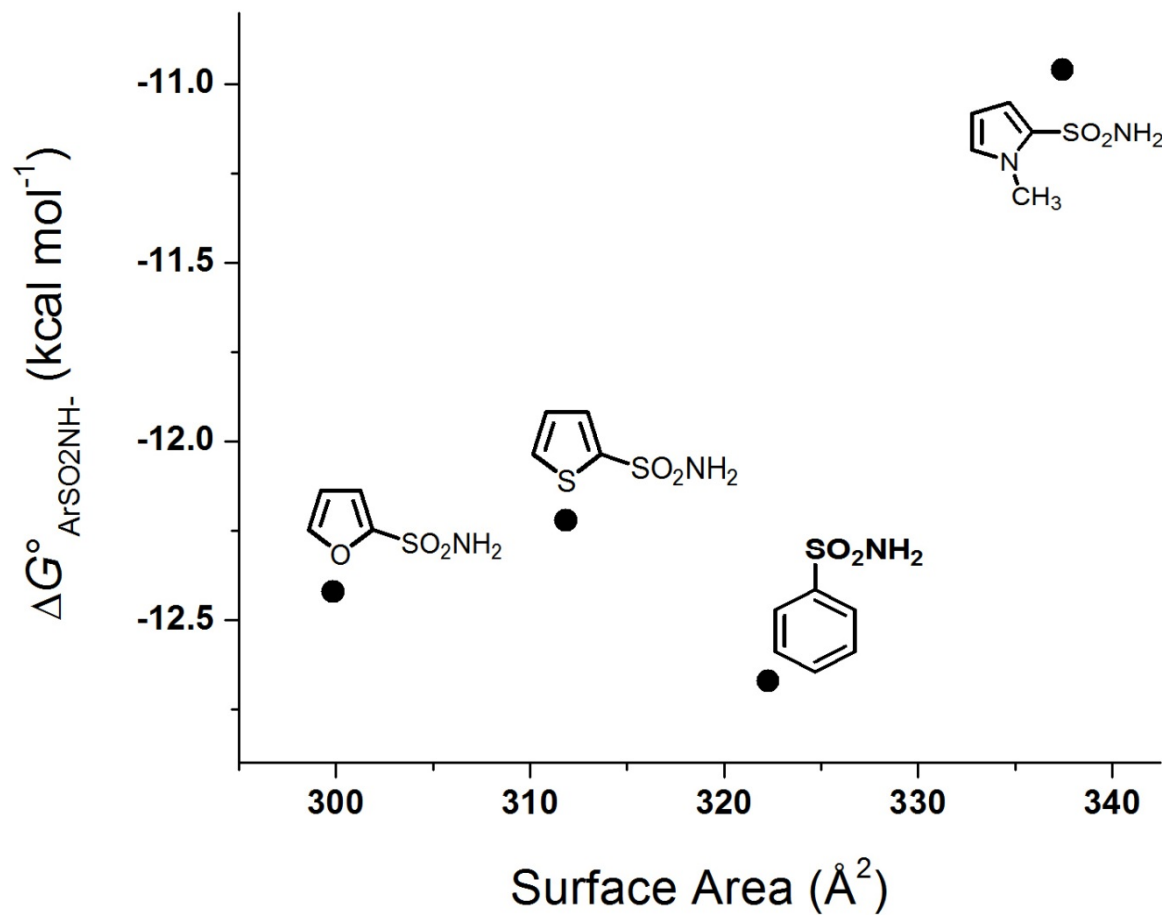
Bovine carbonic anhydrase II

PDB: 1IF4 (adapted)

Constant $-\text{Zn}^{+2}----\text{H} \quad \text{N}-\overset{\overset{\displaystyle O}{\|}}{\underset{\underset{\displaystyle O}{\|}}{S}}-$ R Variable
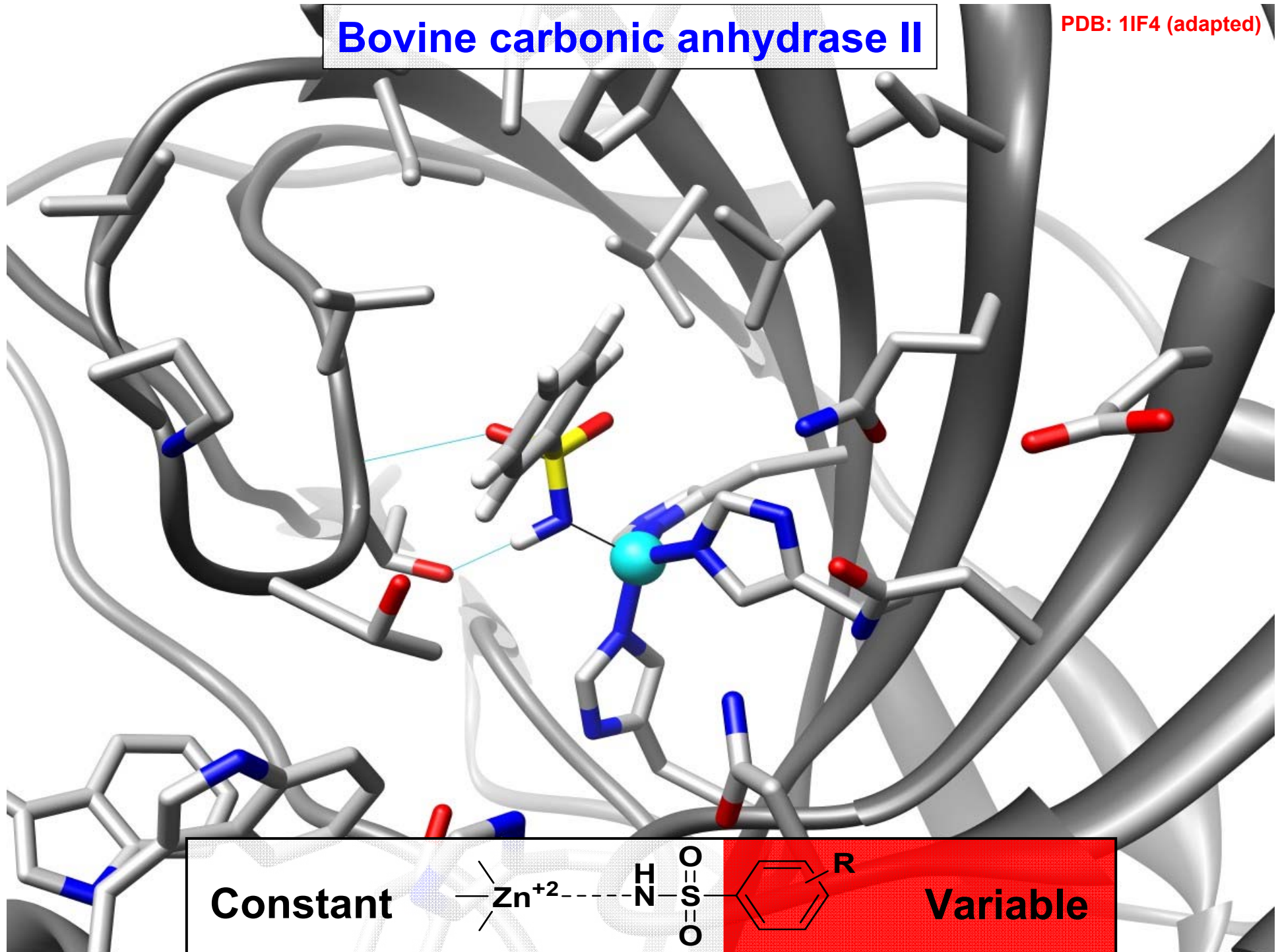
PDB: 1IF4 (adapted)

Polar | Hydrophobic

Monocyclic aromatics have similar K_d
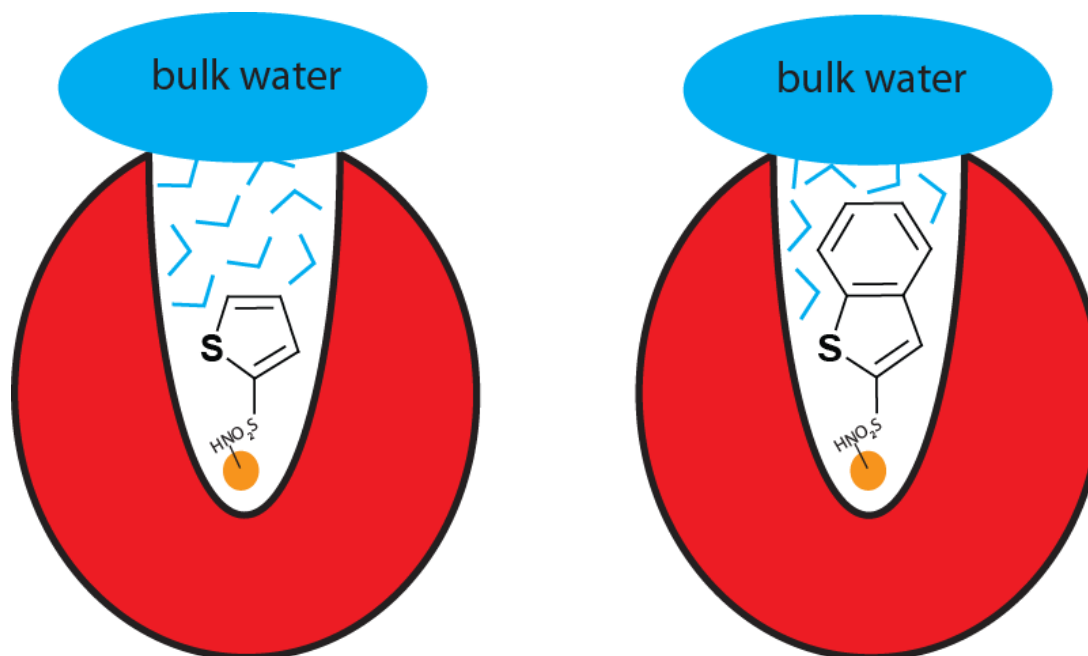

These small differences are not easily interpretable

The "Benzo Addition" Strategy

- **Classical physical-organic perturbation method: examine a common modification that extends hydrophobic surface area**

"benzo" extension

$$H_2NO_2S \longrightarrow H_2NO_2S$$

X = CH, N
Y = O, NH, S
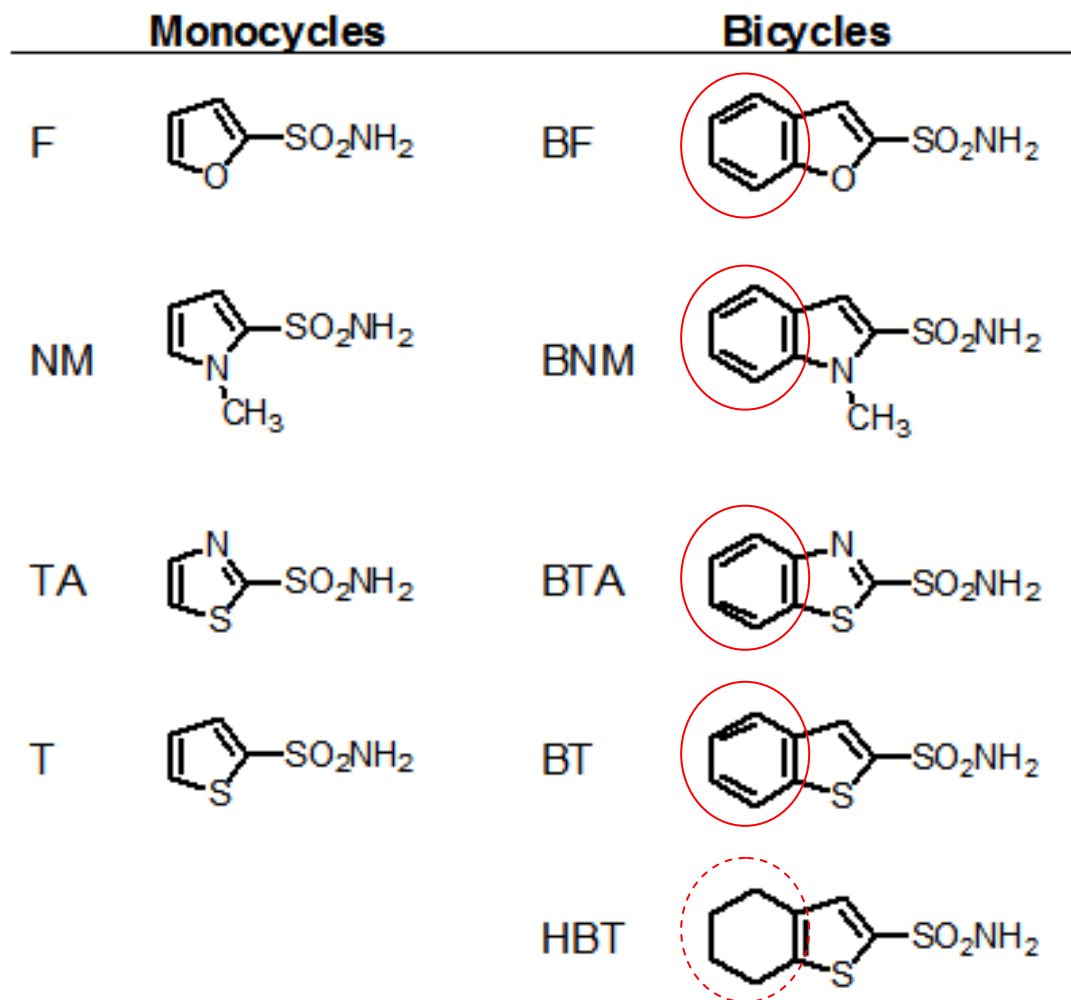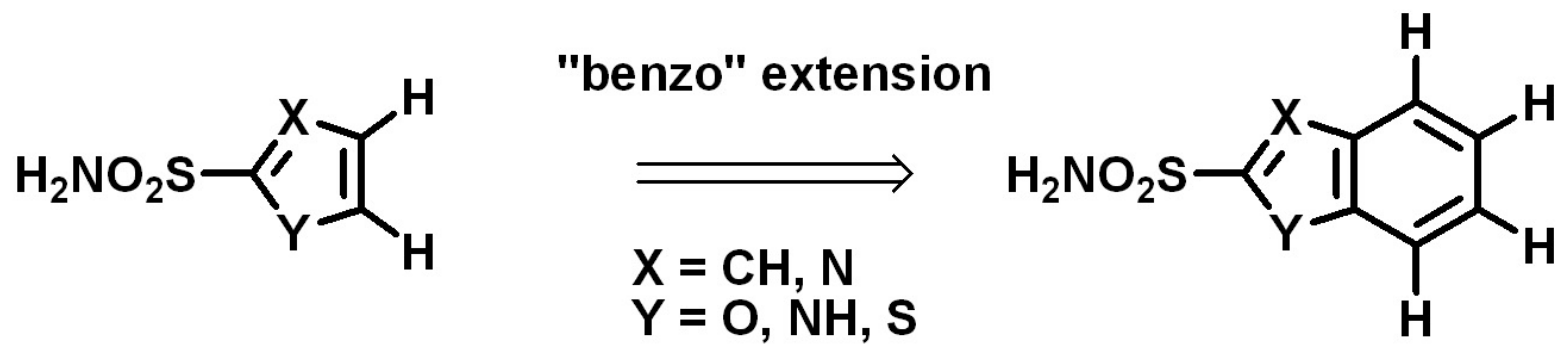
- **Increase the hydrophobic surface area of the ligand (benzo-extension)**
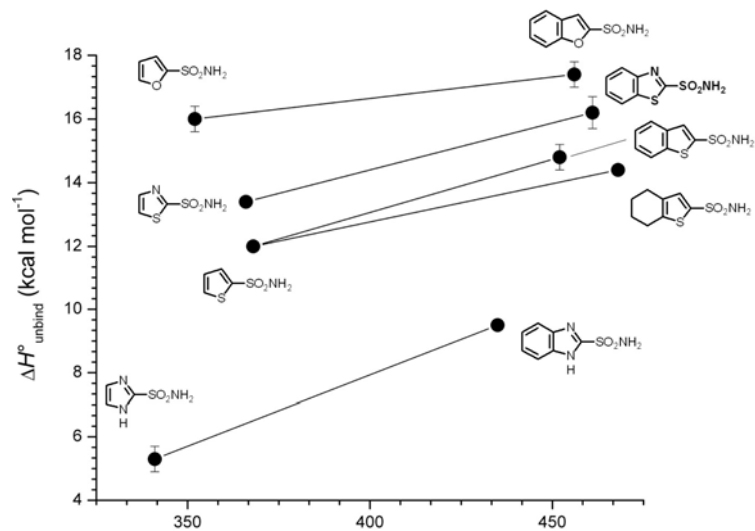
"Benzo Addition:" Summary

- $\Delta\Delta G^0 \sim -20$ to -25 cal/A^2
- Dominated by *enthalpy*, not entropy
- Independent of aromatic/aliphatic. *No "non-classical hydrophobic effect"*

- Requires: a non-Kauzmann interpretation of the hydrophobic effect

DISSOCIATION
protein bound --> buffer

	Monocycles		Bicycles
F	⌬—SO₂NH₂ (furan)	BF	⌬—SO₂NH₂ (benzofuran)
NM	⌬—SO₂NH₂ (N-methylpyrrole)	BNM	⌬—SO₂NH₂ (N-methylindole)
TA	⌬—SO₂NH₂ (thiazole)	BTA	⌬—SO₂NH₂ (benzothiazole)
T	⌬—SO₂NH₂ (thiophene)	BT	⌬—SO₂NH₂ (benzothiophene)
		HBT	⌬—SO₂NH₂ (tetrahydrobenzothiophene)
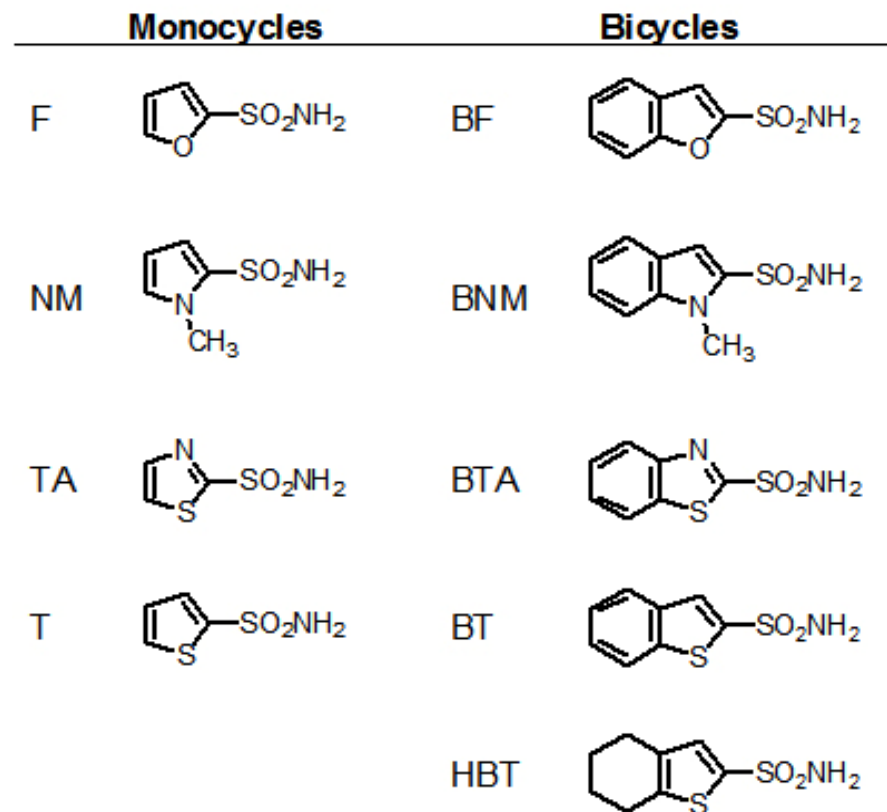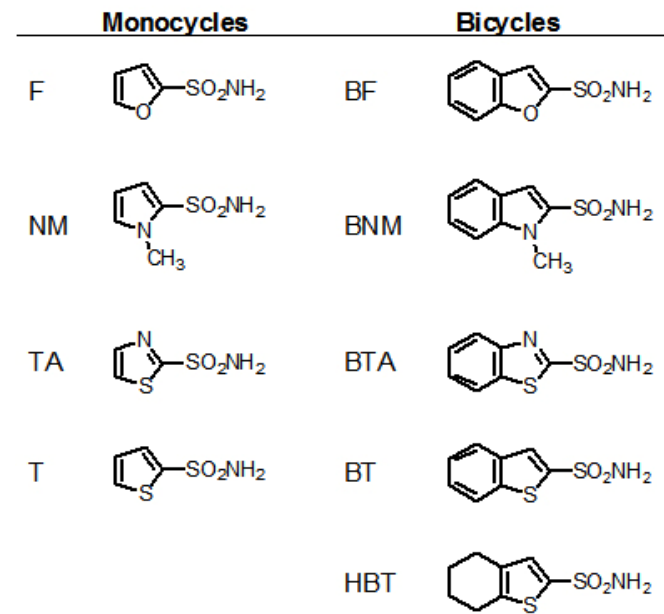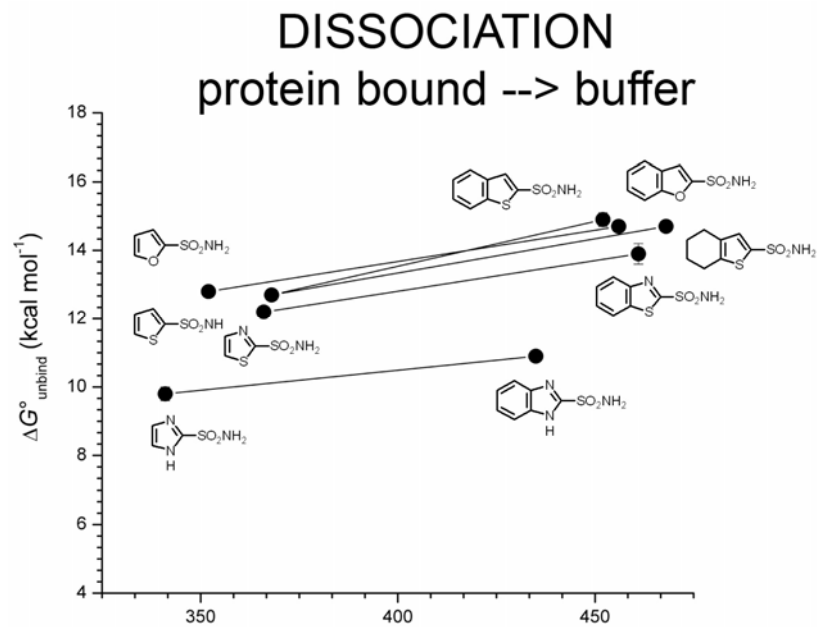
KT predict free energy, but enthalpy and entropy make the opposite contribution
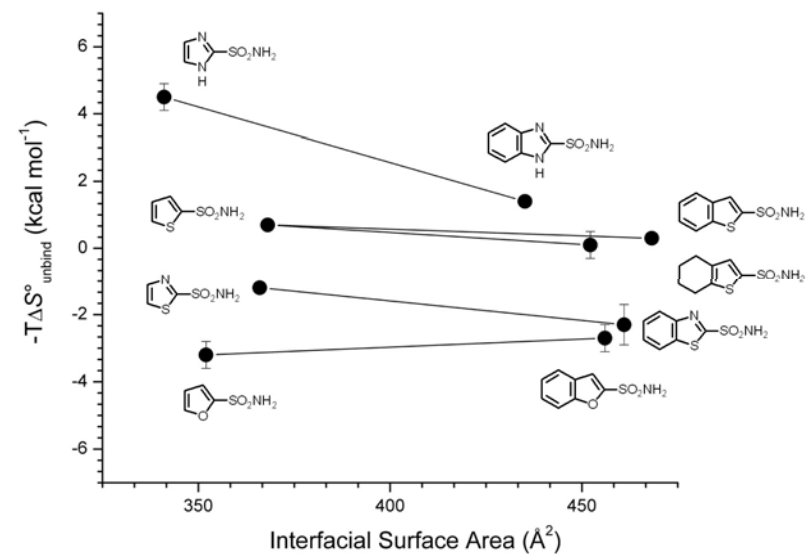
Heat Capacity: Compatible with Hydrophobic Effect
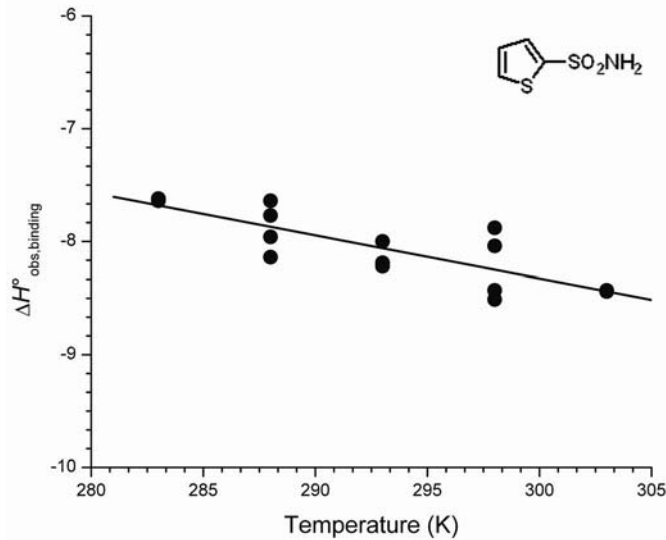
$$\Delta C^\circ_p = -38 \pm 7 \text{ cal mol}^{-1}\text{ K}^{-1}$$
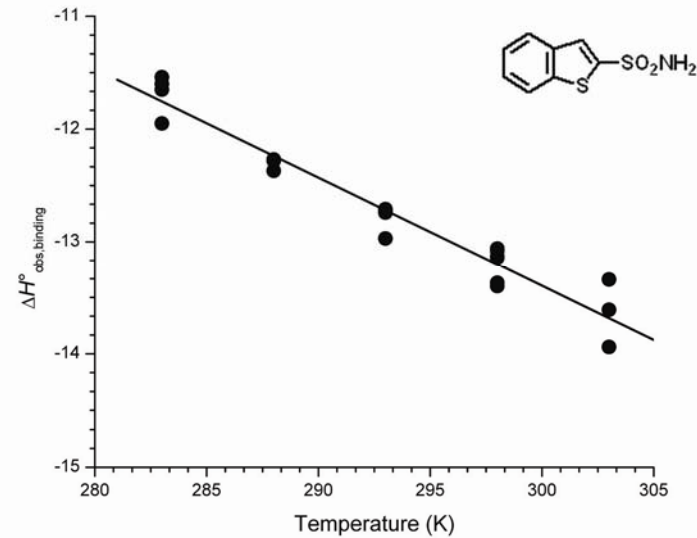
$$\Delta C^\circ_p = -96 \pm 6 \text{ cal mol}^{-1}\text{ K}^{-1}$$

- Negative values of ΔC_p° => hydrophobic effect/solvation changes
- ΔC_p° is complicated
  - Opposite contributions from dehydration of polar and nonpolar (Privalov)
  - Protein structure (Sturtevant)
 - Burying water molecules (Connelly, Ladbury)
  - Releasing water molecules

- Ligand Partitioning
- Free energy is the same as binding
- Entropy and Enthalpy are opposite to binding
- KT predicts partitioning

PARTITIONING
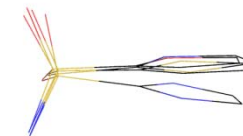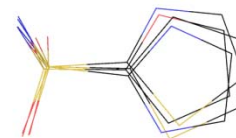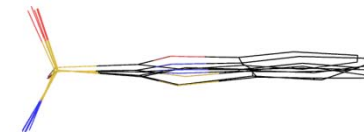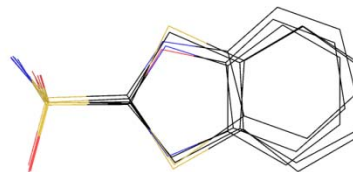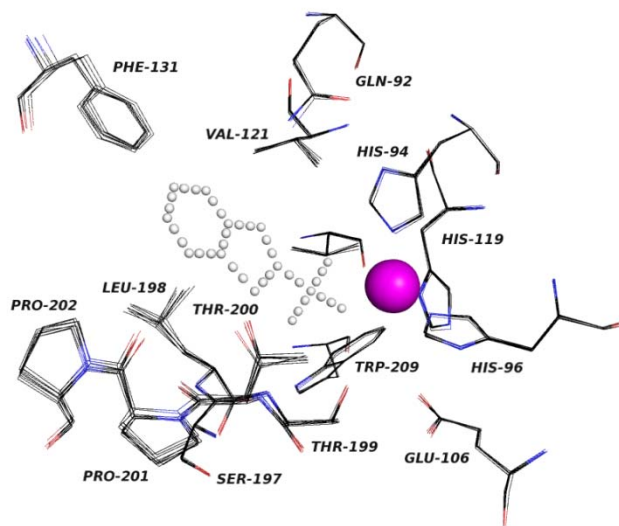OctOH --> buffer

Accessible Surface Area (Å^2)
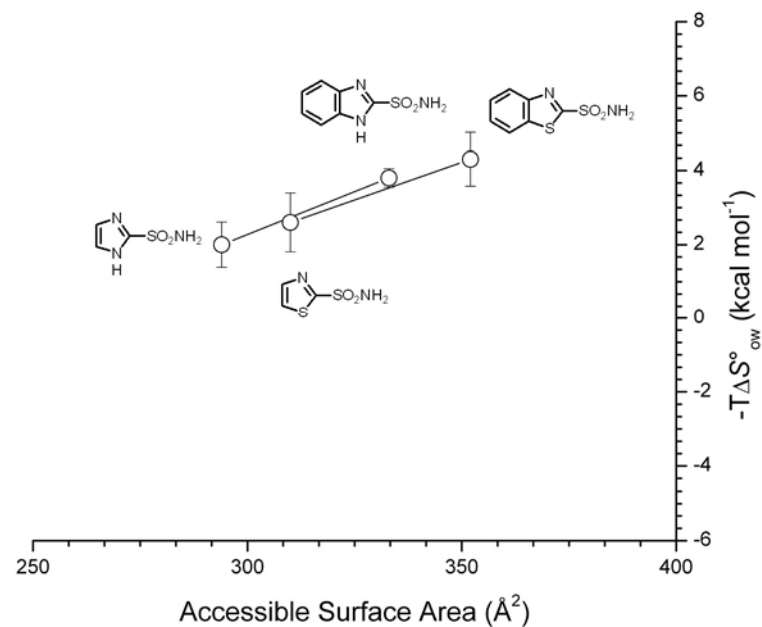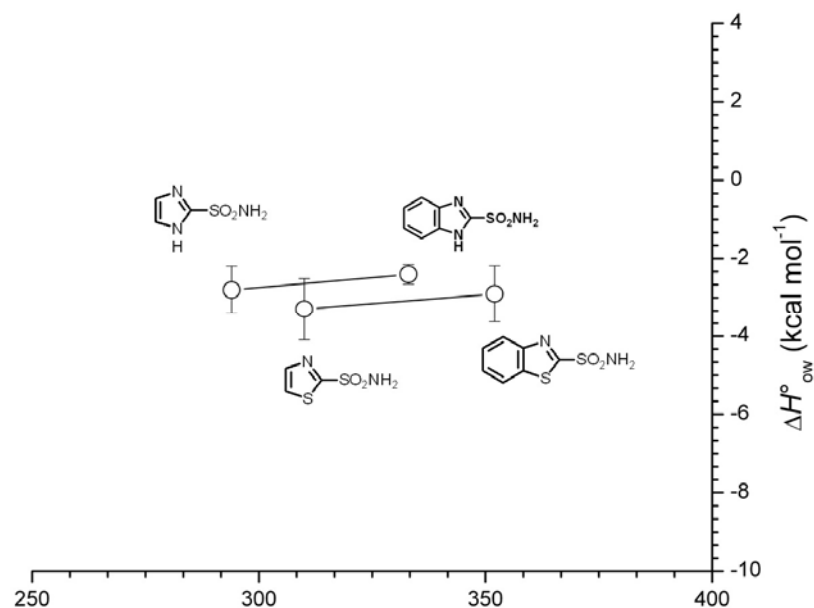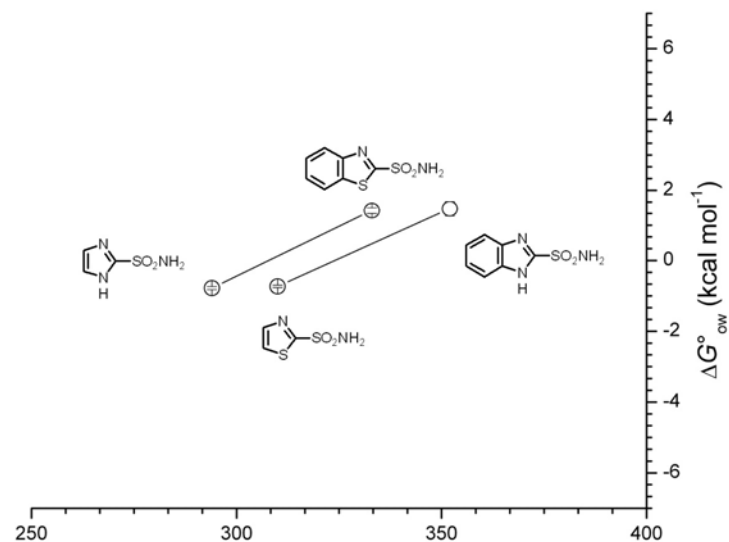
- "Hydrophobicity " of aryl sulfonamide ligands (buffer → octanol) and "hydrohobicity "(buffer → Protein active site) have the same value in free energy, but completely different mechanisms

- …implies many "hydrophobic effects"?
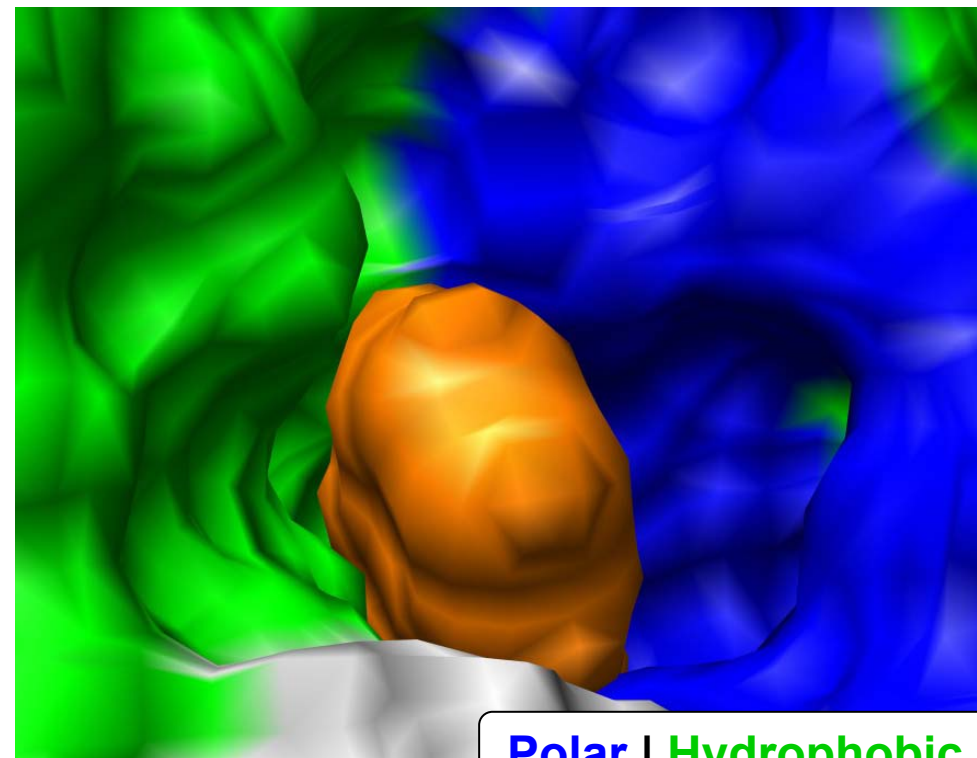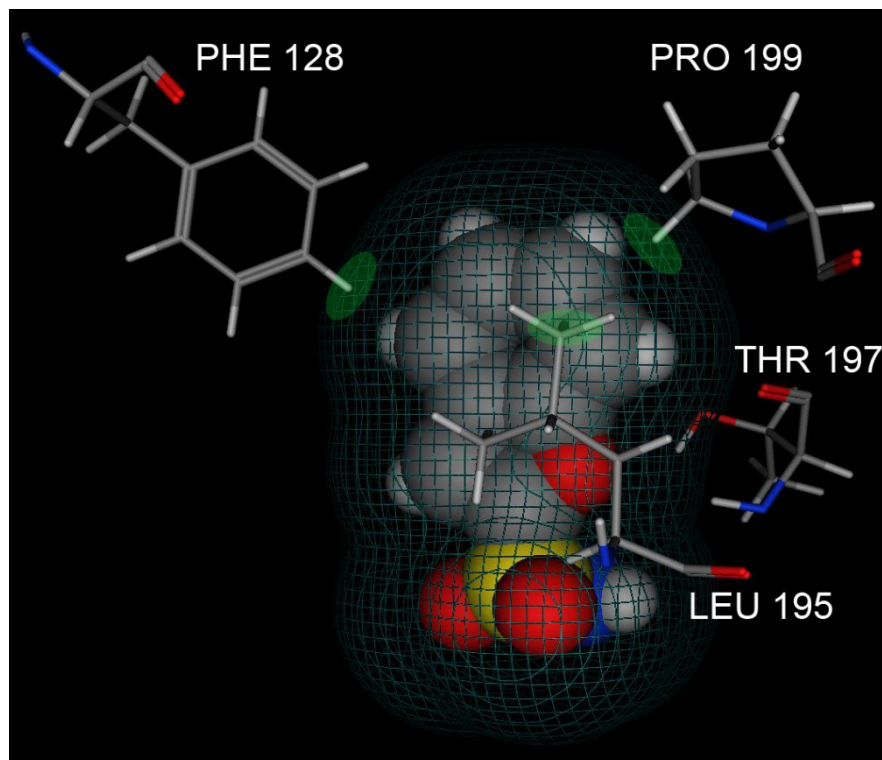
- Entropy/enthalpy compensation?

.
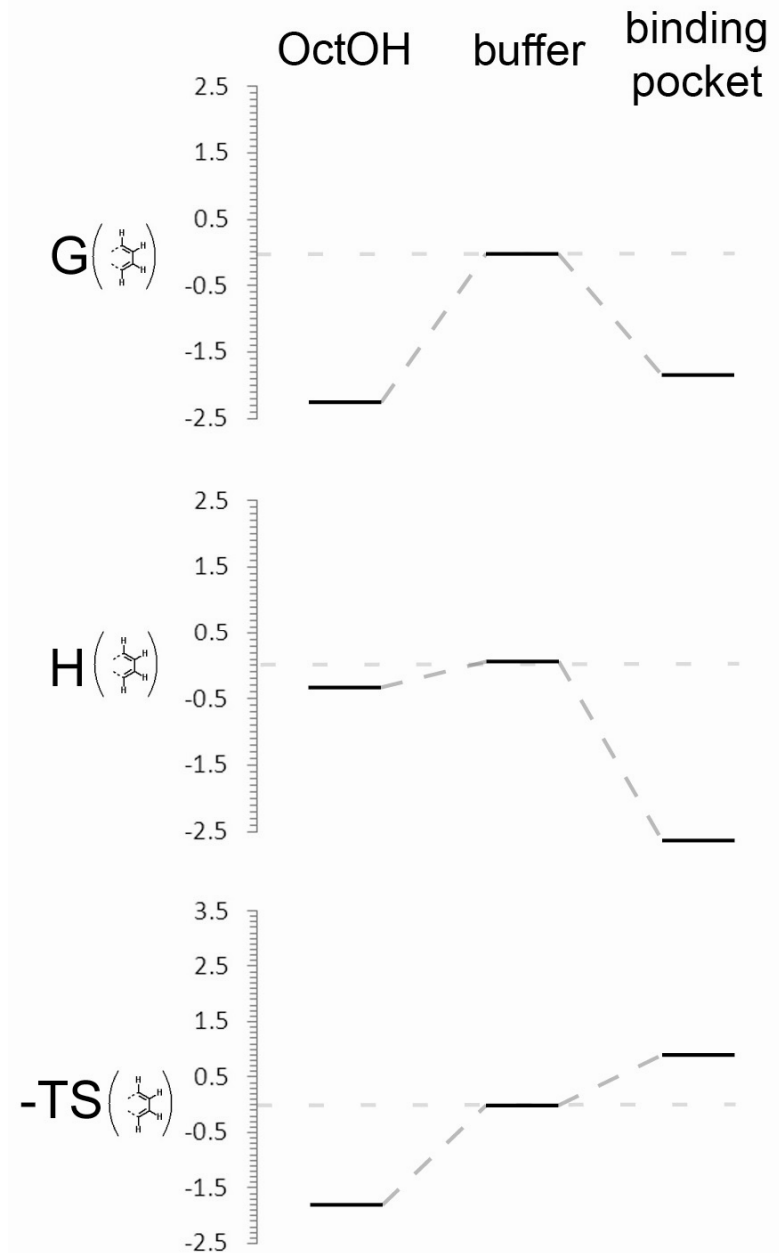
Structure of the protein and the ligands is invariant
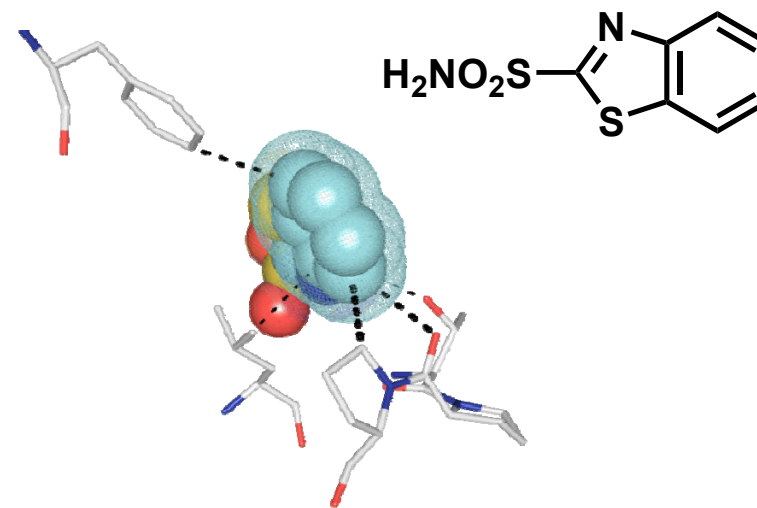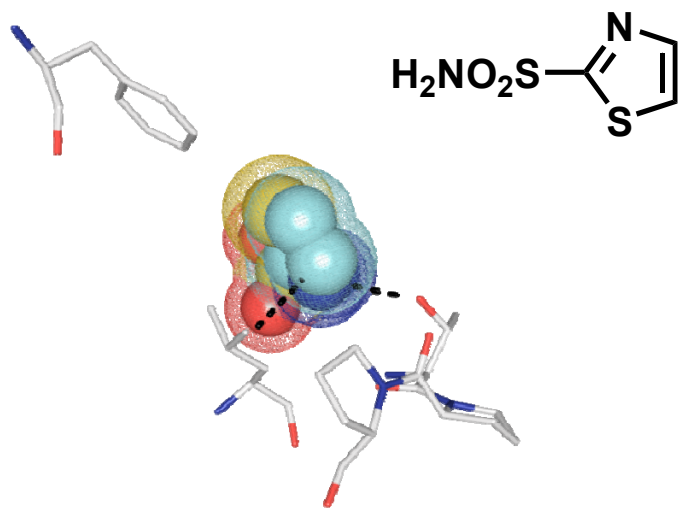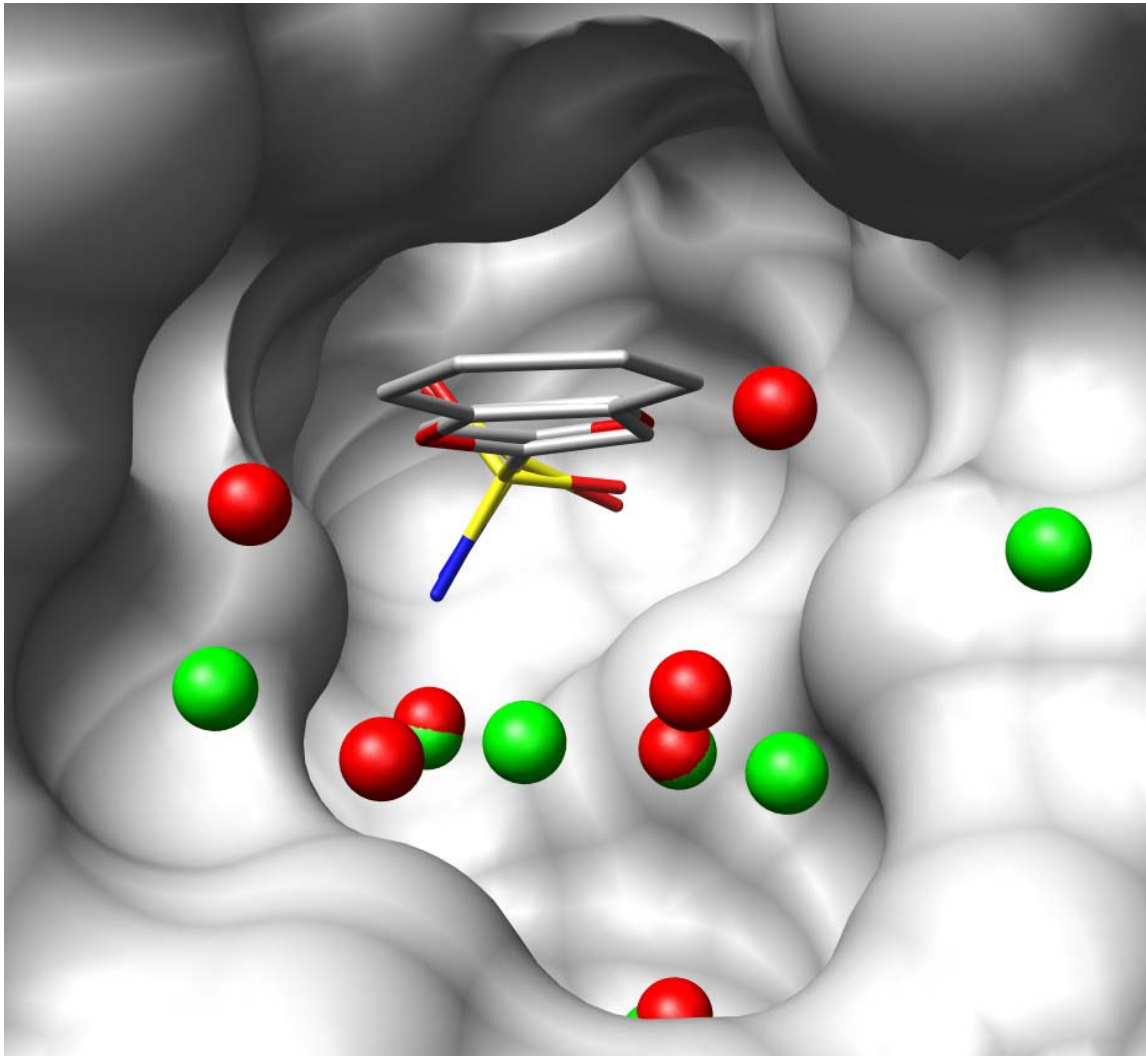

Few contacts between benzo and hydrophobic wall

Fused benzo ring makes few contacts with protein

1.3 Å

PHE 128   PRO 199

THR 197

LEU 195

Polar | Hydrophobic

Minimal contacts between fused ring and protein
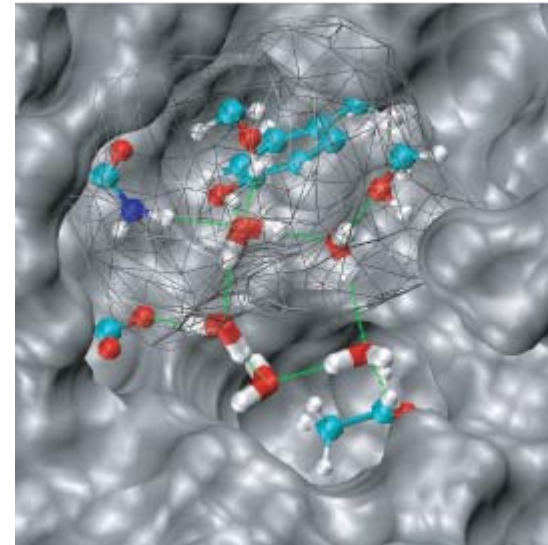
Most of the fused ring is solvent-exposed

Ligand is highly solvent-exposed

Fused ring appears to affect crystallographic water positions

Hypothesis: observed thermodynamic effect for "hydrophobic binding" is solvent-mediated

- **Water in cavities**

  - **Rossky, Lazaridis, Berne, Friesner**

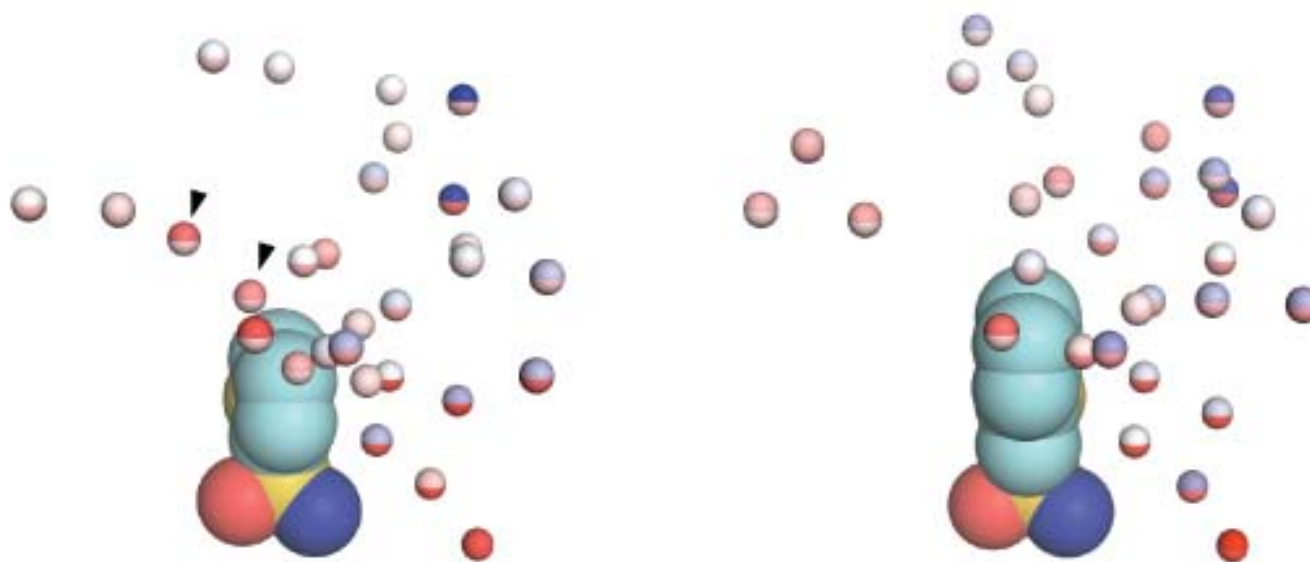 - **Free energy depends on the geometry of the cavity, polarity of the surface**

- **WaterMap ® (Schrodinger)**

  - **explicit water MD**

  - **estimates ΔH and ΔS for water**
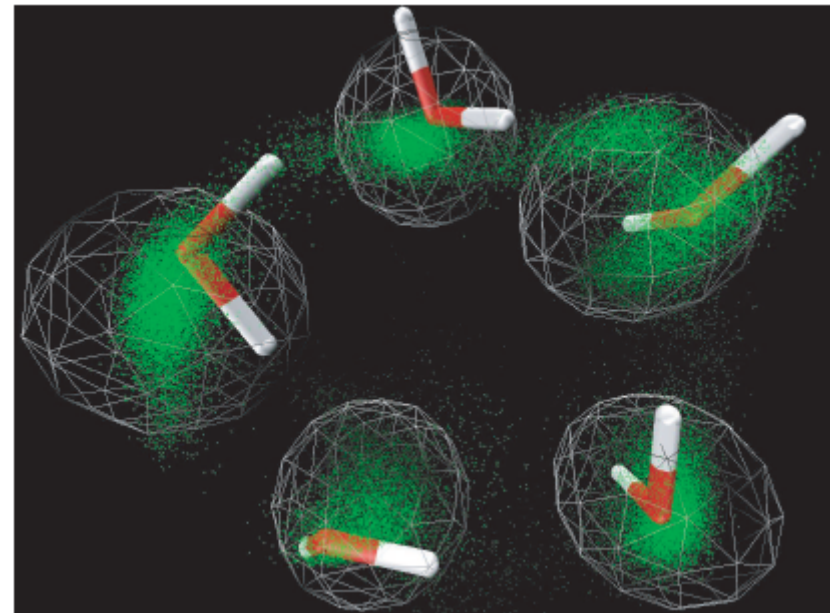
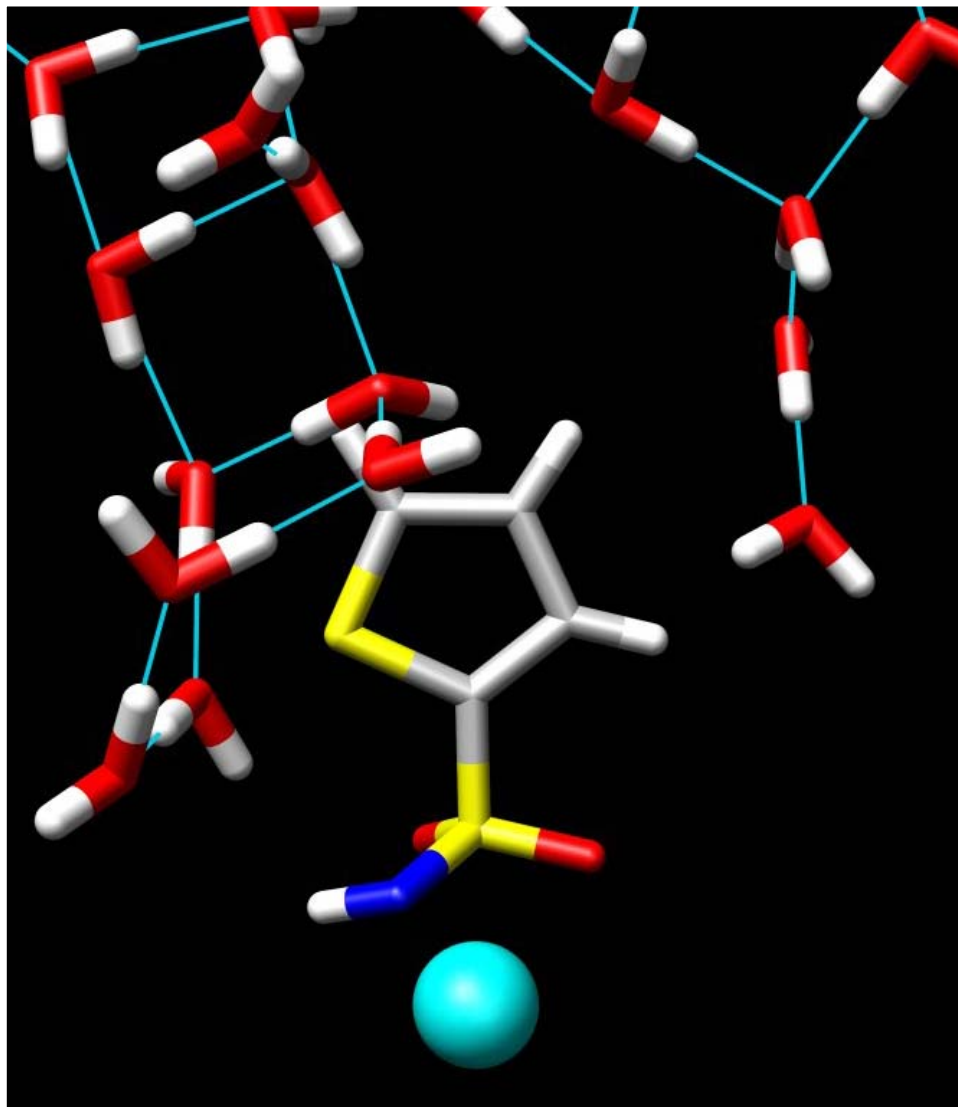Young, Abel, Kim, Berne, Friesner *PNAS* **2007**

WaterMap predicts $-T\Delta\Delta S^{\circ} \approx 0$ cal mol^{-1}
$\Delta\Delta H^{\circ} \approx -3$ cal mol^{-1}

ΔH°_{WM}

$-T\Delta S^{\circ}_{WM}$

+4 kcal mol^{-1} -4 kcal mol^{-1}

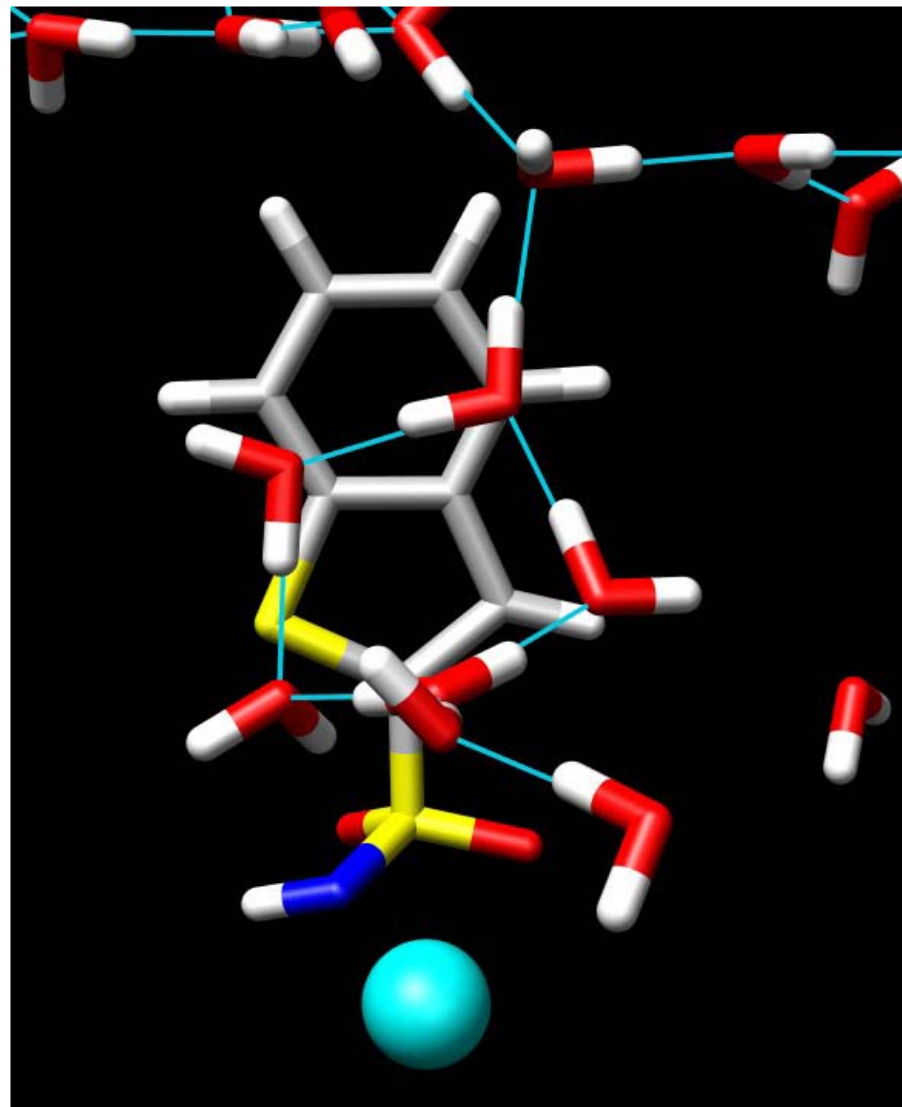unfavorable   favorable

Indistinguishable from thermochemical measurements

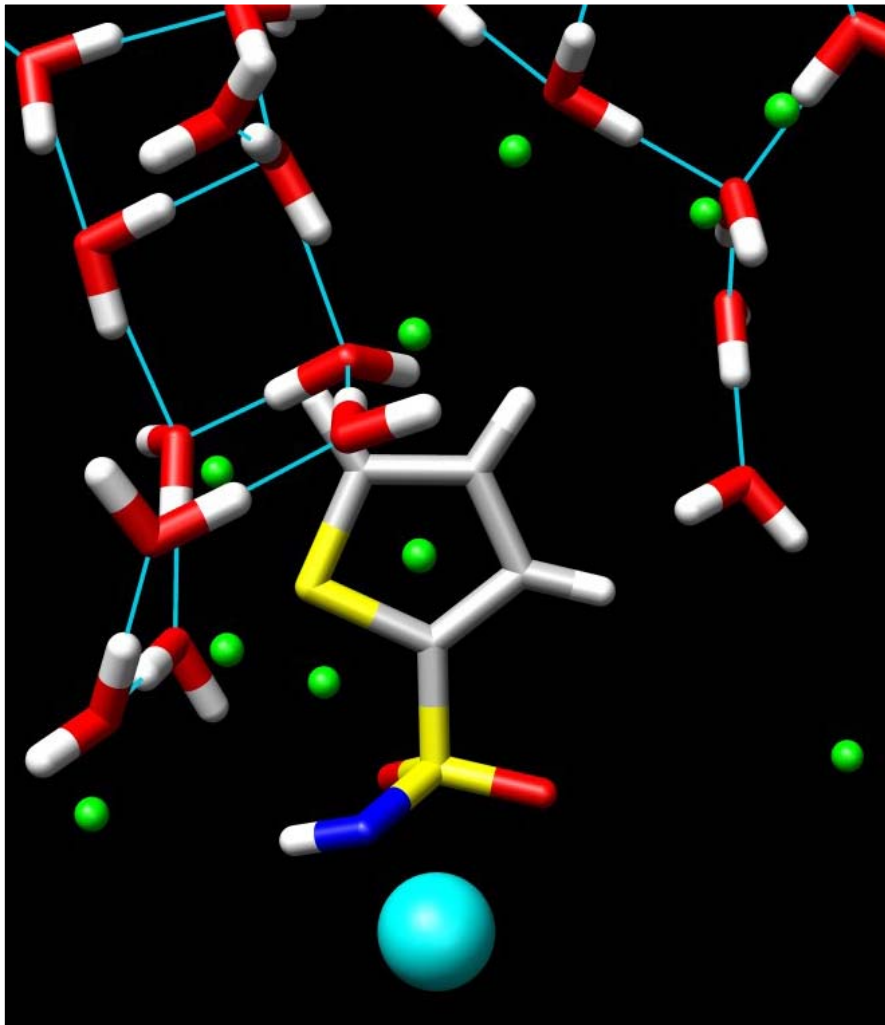# Comparison to crystallographic waters?

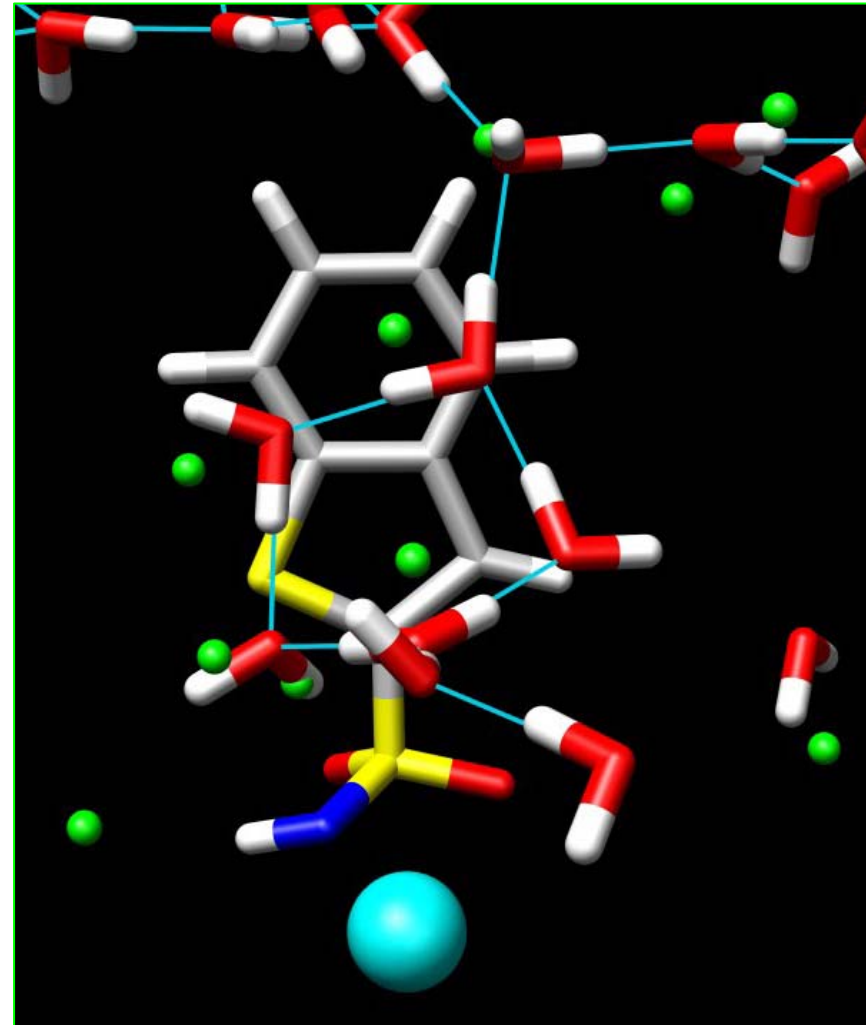Thiophenesulfonamide

Benzthiophenesulfonamide

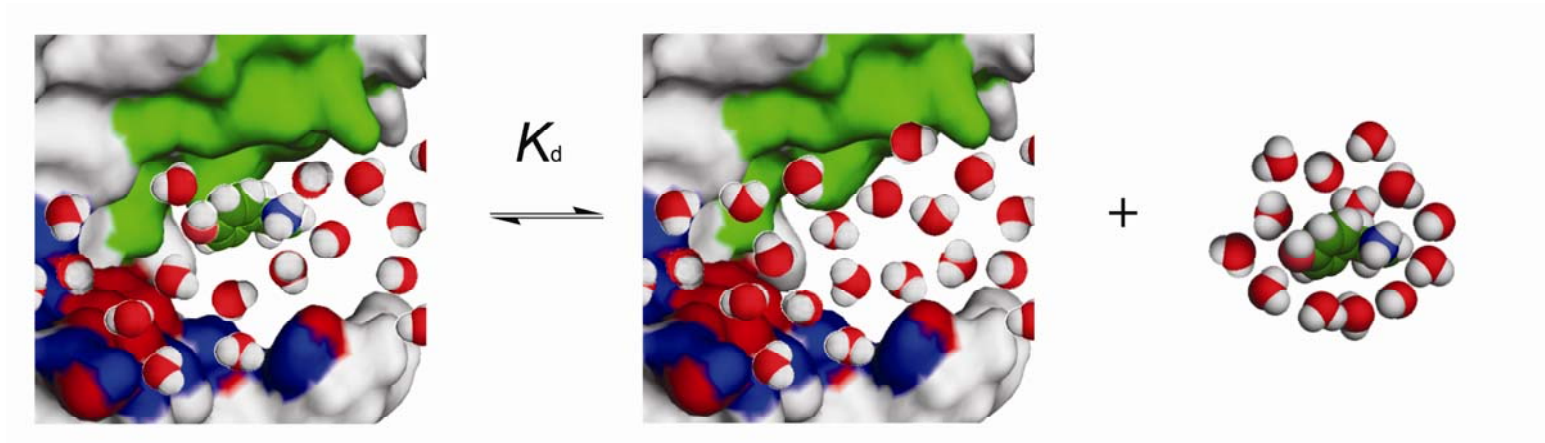# Crystallographic waters predicted by modeling

Thiophenesulfonamide

Benzthiophenesulfonamide

So: What *is* the Hydrophobic

- It is *not* the apposition of two non-polar surfaces with release of hydrogen-bond networks (e.g., not "lock and key" or KT)

- It might be:
  - The "shape of the water" in the binding pocket, rather than the shape of the pocket
  - The displacement of energetically unfavorable water into bulk water from active site, and release of surface water from ligand

What is Needed?

- More examples coupling structure and thermodynamics.
- For computation: better (or better justified) potential functions (for H), and much faster computation (for S)
- ITC that is more routine, and requires less protein.
- Better fundamental understanding of water and hydrophobicity
- Tests based on *protein* structure: mutagenesis
- A sound theory of molecular recognition in water

Maximal Affinity for the Binding of Small Molecule Ligands
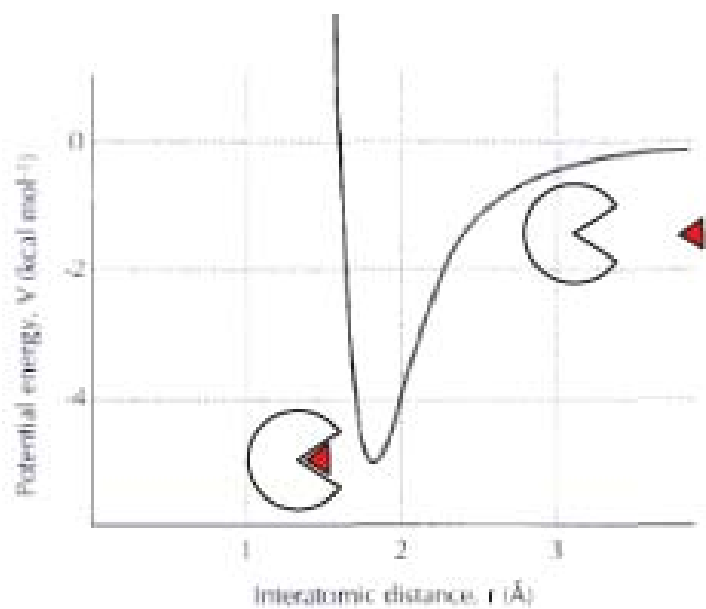
13 macromolecules interacting with
136 different ligands
186 different combinations examined
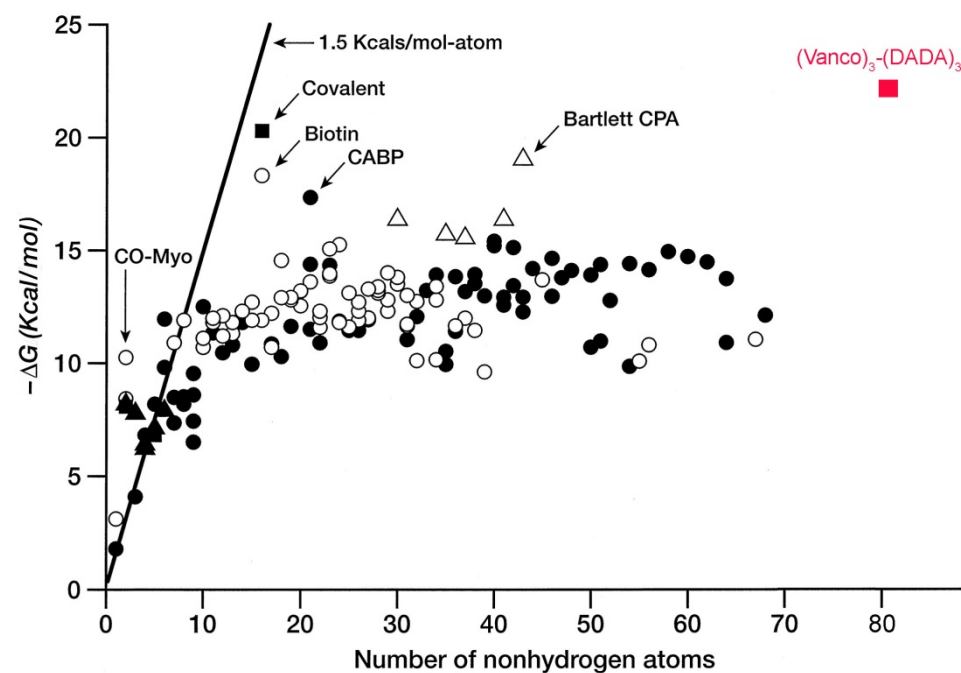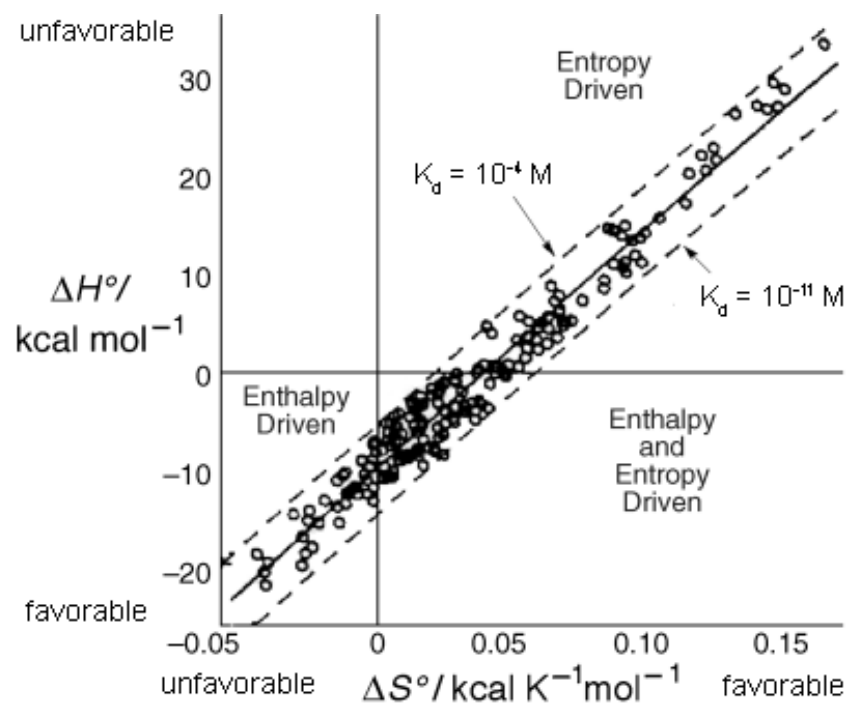
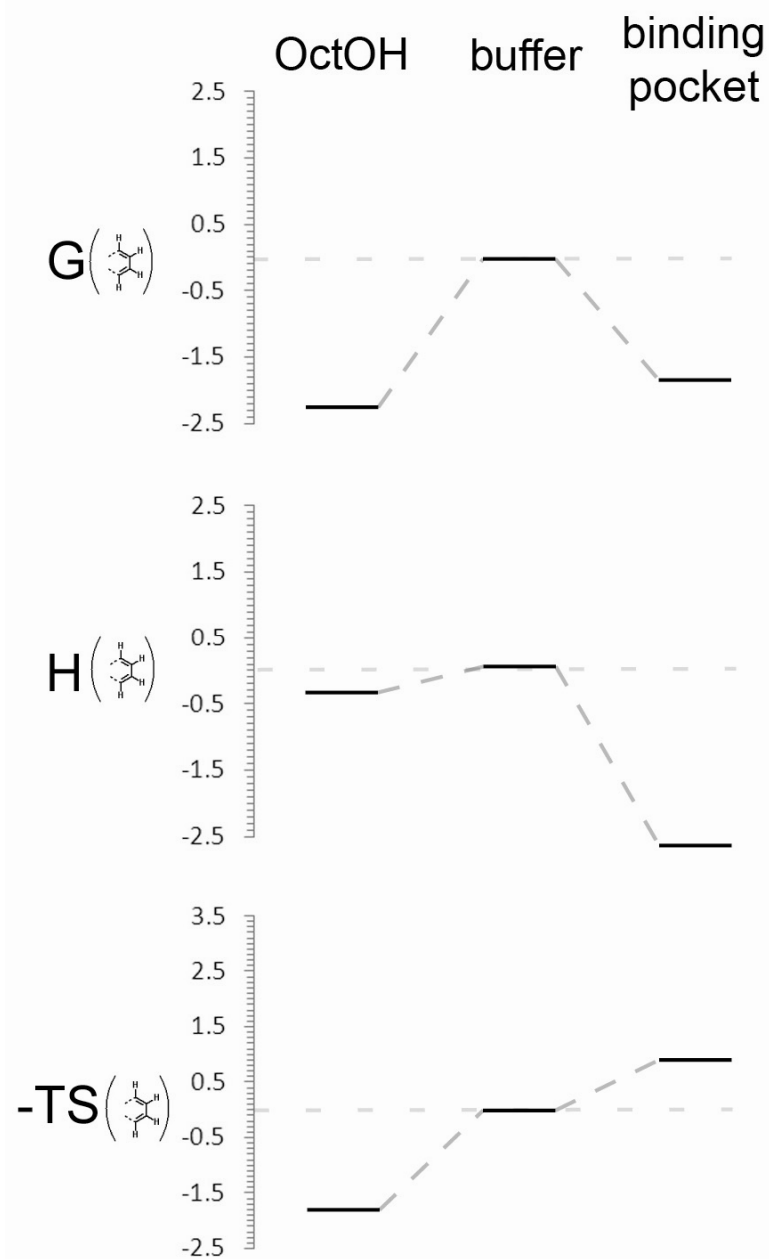Gilli, P. et al. *J. Phys. Chem.* **1994**, *98*, 1515-1518.

Kuntz, I.D. et al. *PNAS* **1999**, *96*, 9997-10002.

- Entropy-Enthalpy compensation
- We still don't understand
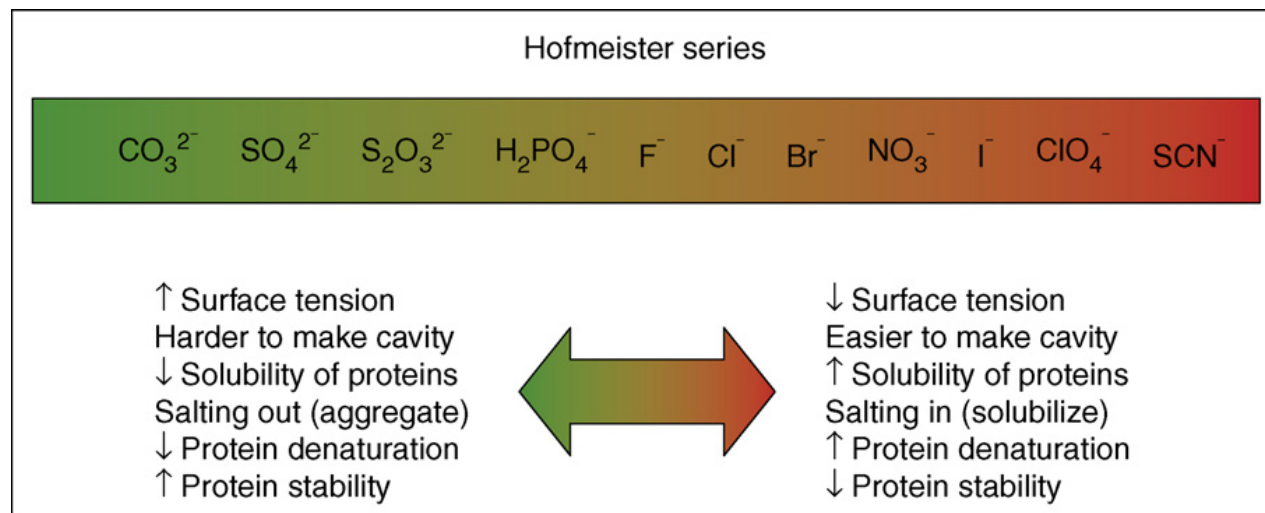

OctOH   buffer   binding pocket

Dunitz, J. *Chem. Biol.* **1995**, *2*, 709-712.
Williams, D.H. and co-workers *Angew. Chem. Int. Ed.* **2004**, *43*, 6596-6616.
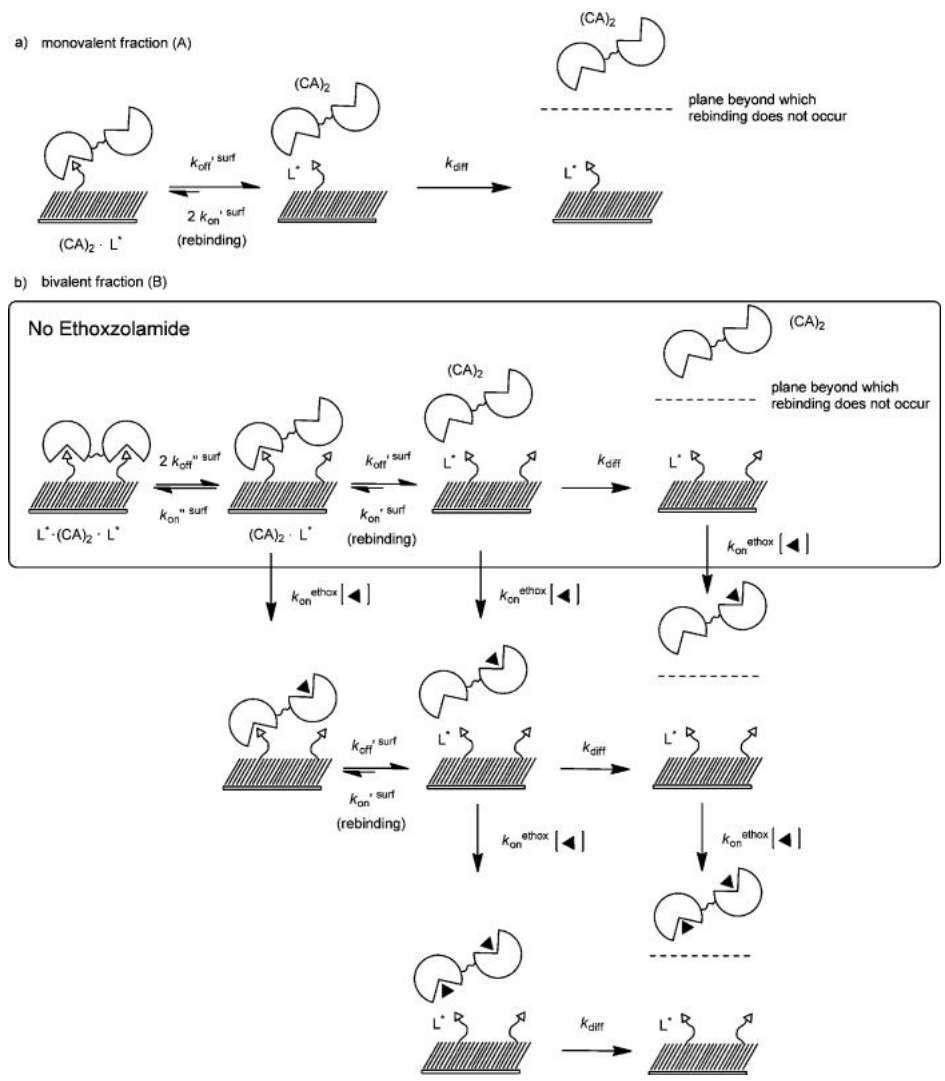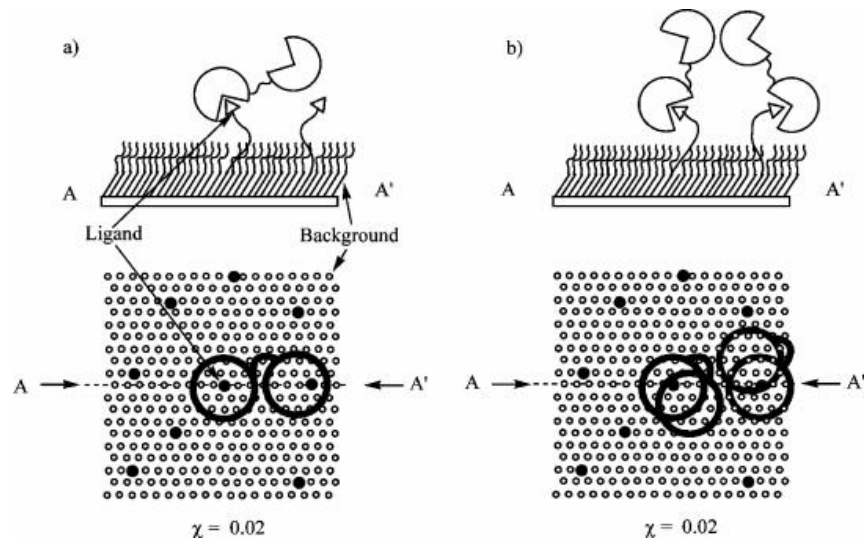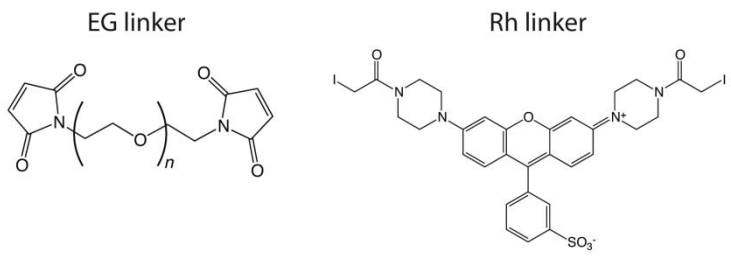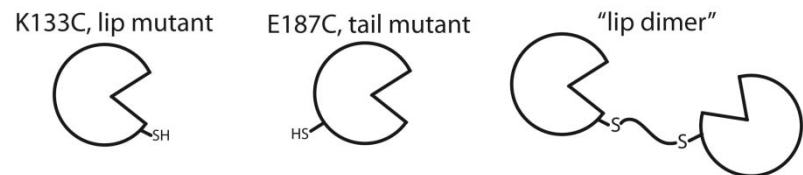
The effect of medium

- $\Delta\Delta G$ for benzo-group independent of medium

- $\triangle G$, $\triangle H$ and $-T\Delta S$ independent of anion and cation

- Urea (1M), glycine (1M) and DMSO (10%) do not effect $\triangle H$ and $-T\Delta S$

- Ethanol (10%) or PEG (10%) make $\triangle H$ more favorable by ~ 1 kcal mol^{-1}

  and $-T\Delta S$ less favorable by ~ 1 kcal mol^{-1}

Hofmeister series

CO_3^{2-} SO_4^{2-} $S_2O_3^{2-}$ $H_2PO_4^-$ F^- Cl^- Br^- NO_3^- I^- ClO_4^- SCN^-

↑ Surface tension
Harder to make cavity
↓ Solubility of proteins
Salting out (aggregate)
↓ Protein denaturation
↑ Protein stability

↓ Surface tension
Easier to make cavity
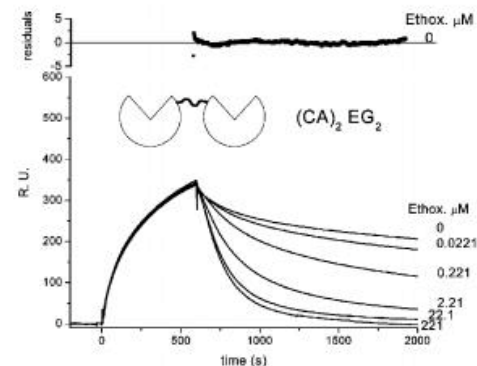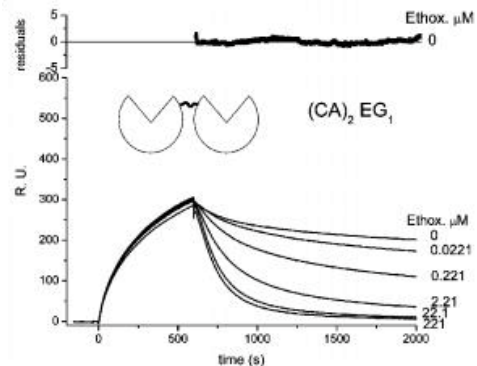↑ Solubility of proteins
Salting in (solubilize)
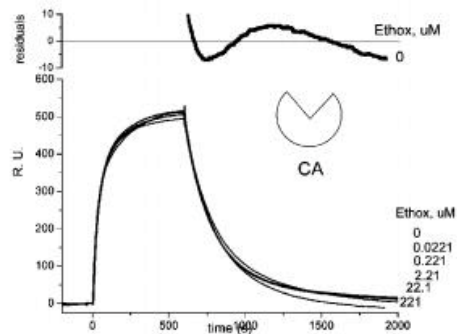↑ Protein denaturation
↓ Protein stability

Zhang, Cremer Curr Opin Chem Biol **2006**, *10*, 658

- **Phil Snyder**
- **Matt Lockett**
- **VJ Krishnamurthy**
- **Demetri Moustakis**

- **Annie Heroux (Argonne National Lab)**
- **Woody Sherman (Schrodinger)**

K133C, lip mutant   E187C, tail mutant   "lip dimer"

EG linker   Rh linker

a) monovalent fraction (A)

$(CA)_2 \cdot L^*$ $(CA)_2$ plane beyond which rebinding does not occur

$k_{off}^{'\,surf}$ k_{diff}

$2\,k_{on}^{'\,surf}$ (rebinding)

b) bivalent fraction (B)

No Ethoxzolamide $(CA)_2$

plane beyond which rebinding does not occur

$L^* \cdot (CA)_2 \cdot L^*$ $(CA)_2 \cdot L^*$

$2\,k_{off}^{''\,surf}$ $k_{off}^{'\,surf}$ k_{diff}

$k_{on}^{''\,surf}$ $k_{on}^{'\,surf}$ (rebinding)

$k_{on}^{ethox}\,[\blacktriangleleft]$ $k_{on}^{ethox}\,[\blacktriangleleft]$ $k_{on}^{ethox}\,[\blacktriangleleft]$

$k_{off}^{'\,surf}$ k_{diff}

$k_{on}^{'\,surf}$ (rebinding)

$k_{on}^{ethox}\,[\blacktriangleleft]$ $k_{on}^{ethox}\,[\blacktriangleleft]$

k_{diff}

a)   Ligand   Background

$\chi = 0.02$

b)

$\chi = 0.02$

	$\Delta G^{o\prime}_{surf}$, kcal mol^{-1}	$K_d^{\prime\,surf}$, nM	$\Delta G^{o\prime}_{avidity,\,surf}$, kcal mol^{-1}	$K_d^{\prime\,avidity,surf}$, nM	enhancement ($K_d^{surf} / K_d^{avidity,surf}$)
HCA II (K133C, lip mutant)	-9.6 ± 0.2	89	-- --	-- --	-- --
HCA II (E187C, tail mutant)	-9.6 ± 0.2	86	-- --	-- --	-- --
$(CA)_2 EG_1$	-9.3 ± 0.1	160	-11.9 ± 0.1	1.8	50 ± 14
$(CA)_2 EG_2$	-9.4 ± 0.2	140	-12.0 ± 0.2	1.7	50 ± 20
$(CA)_2 EG_3$	-9.3 ± 0.1	150	-12.0 ± 0.2	1.5	60 ± 20
$(CA)_2 Rh_{front}$	-9.1 ± 0.1	200	-11.8 ± 0.3	2.2	40 ± 20
$(CA)_2 Rh_{back}$	-9.7 ± 0.1	78	-11.8 ± 0.1	2.3	37 ± 9
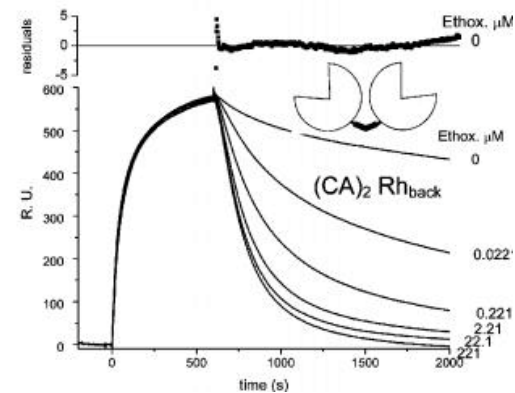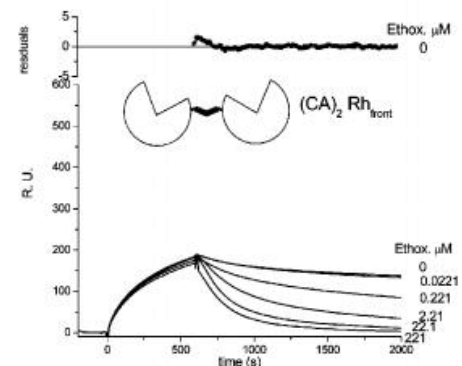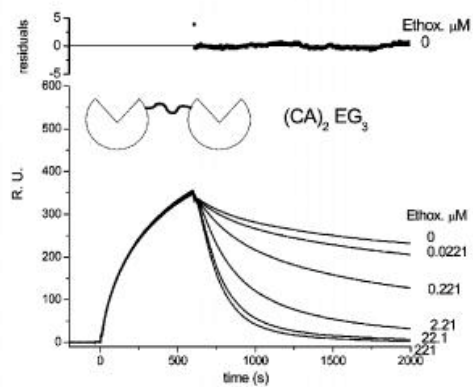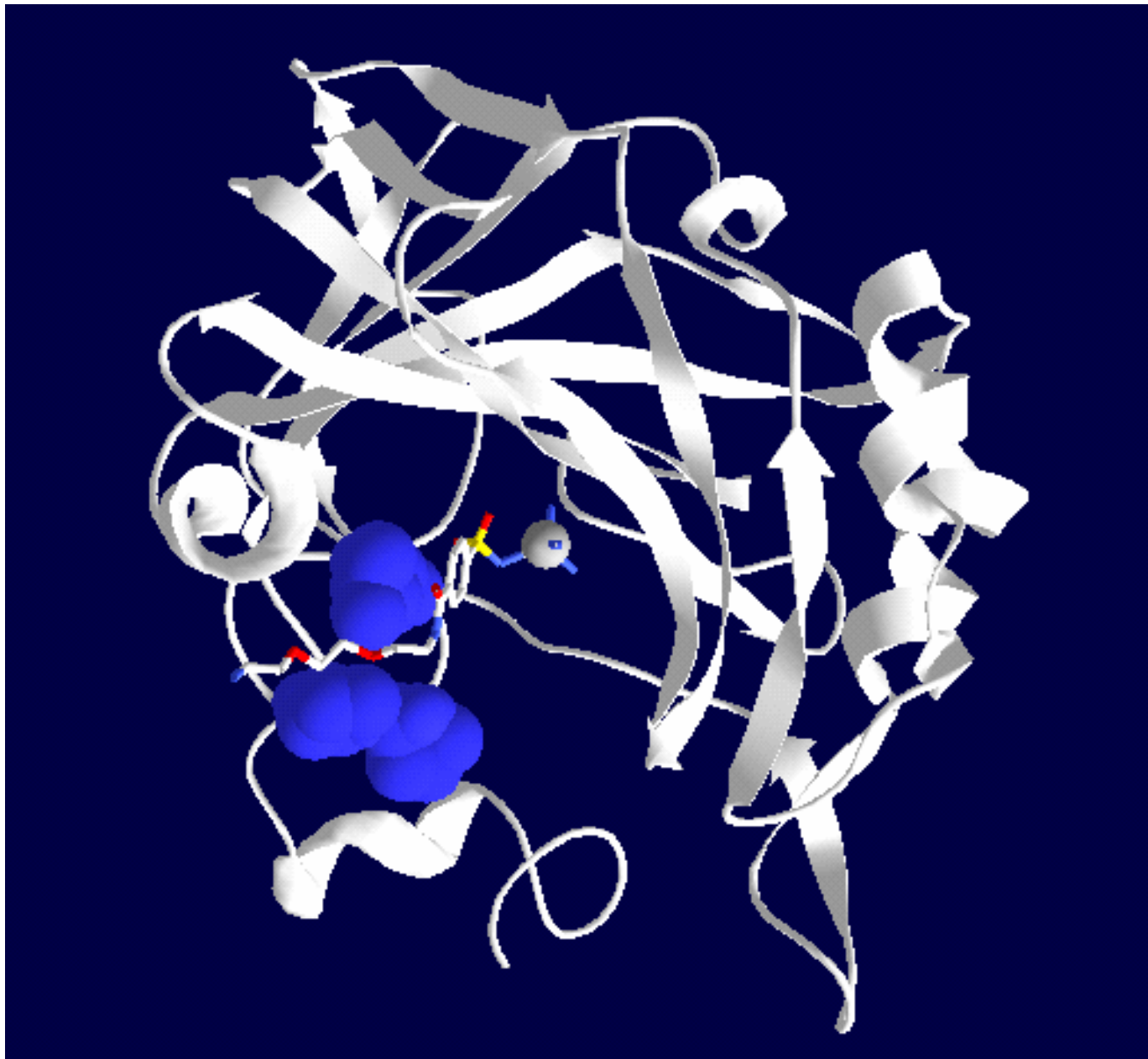
Bovine Carbonic anhydrase II

PDB: 1V9E
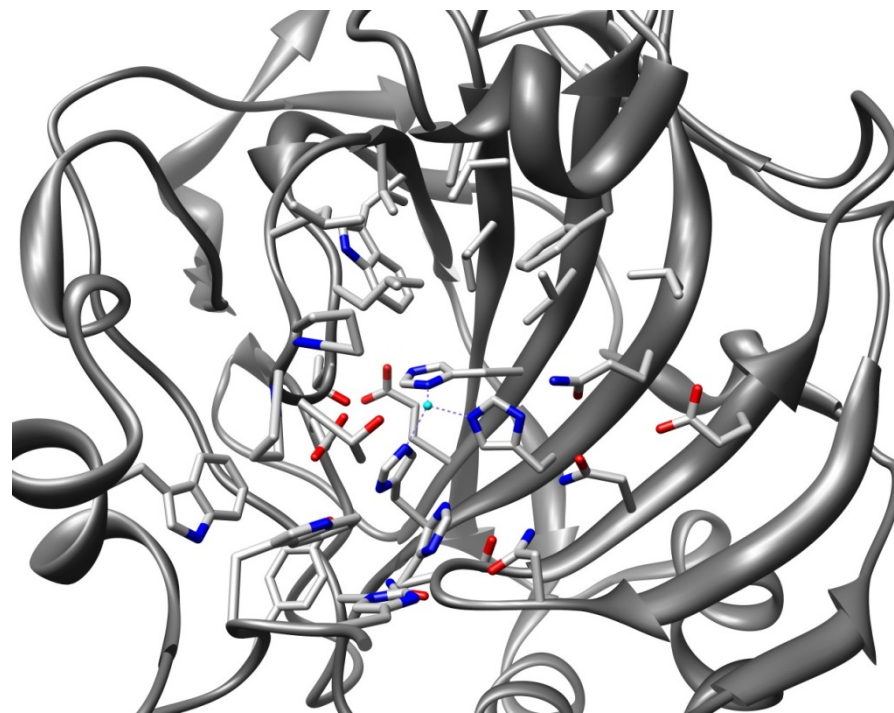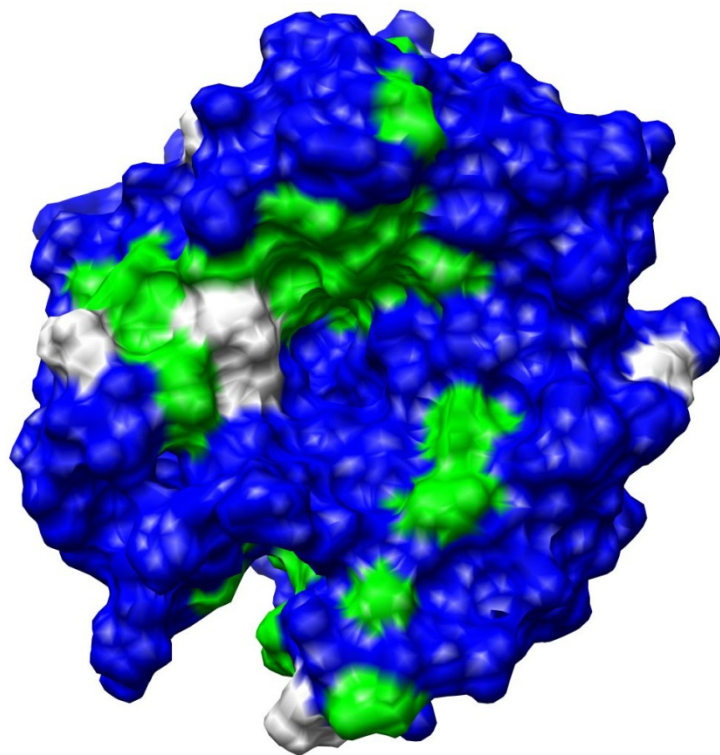
Monocyclic aromatics have similar K_d
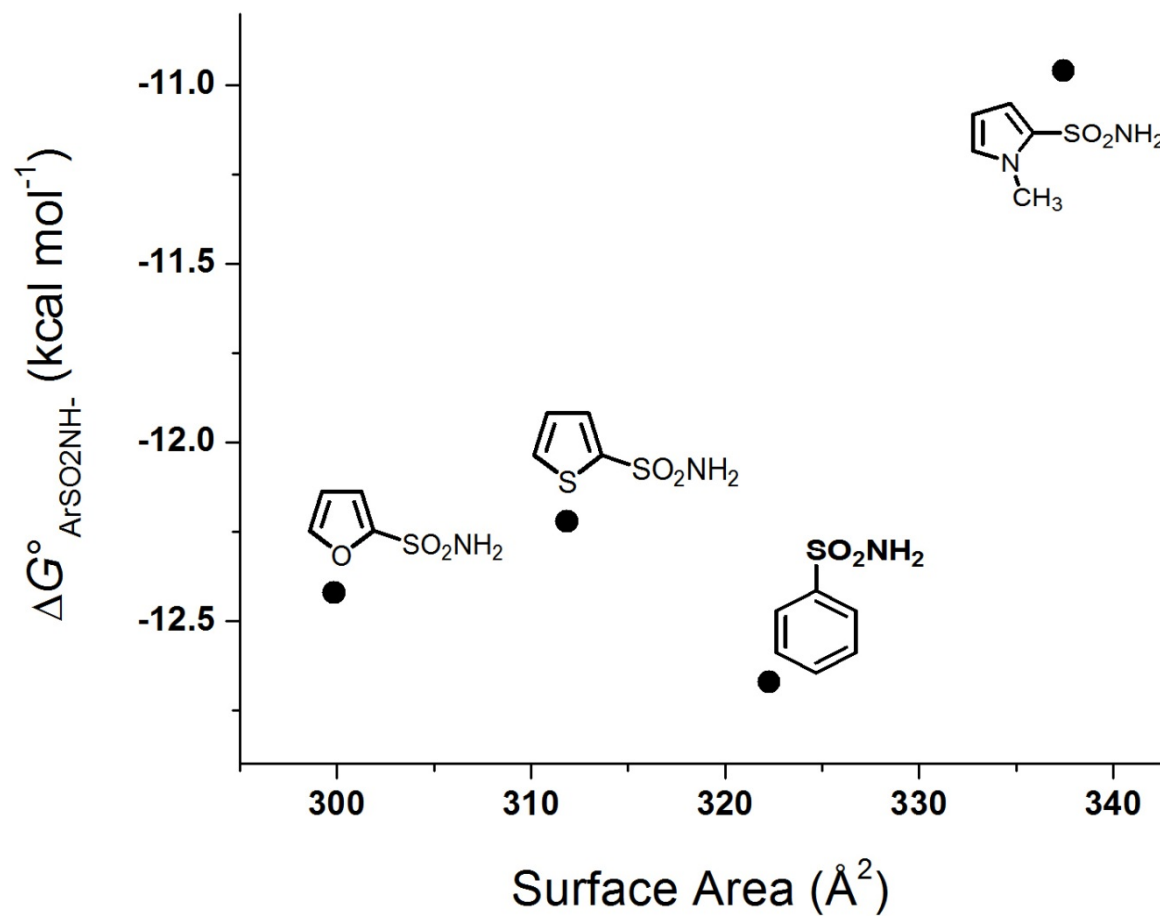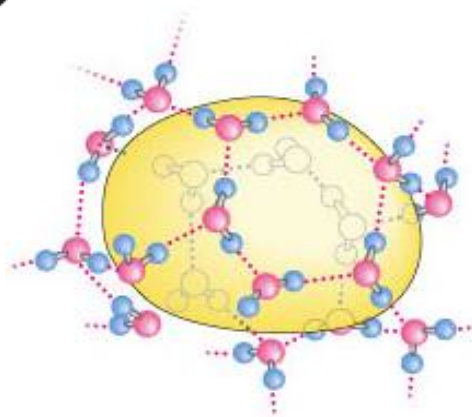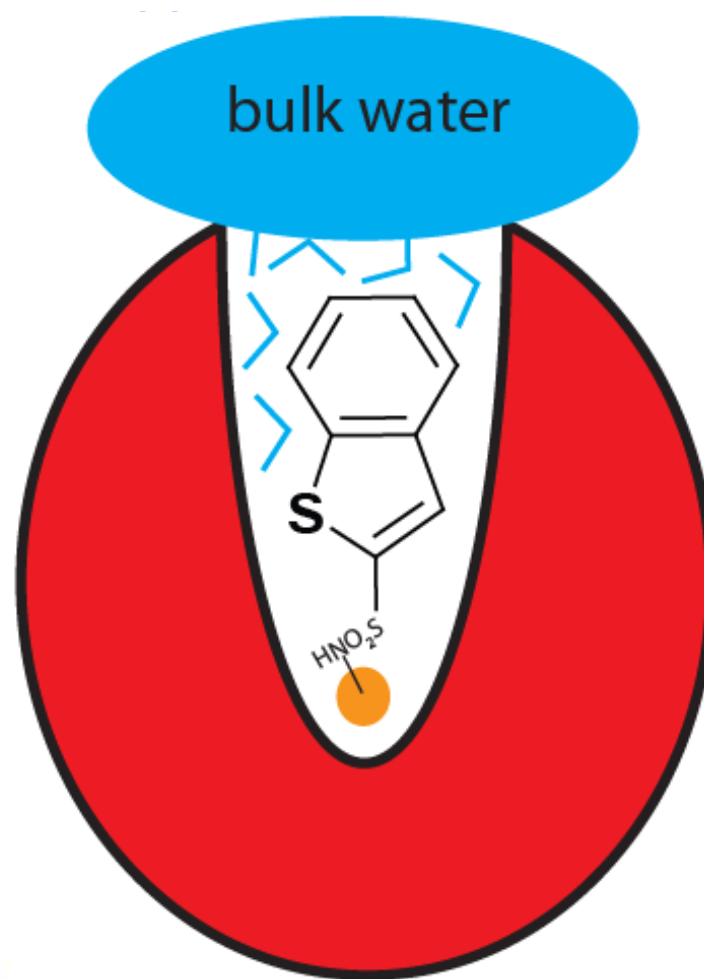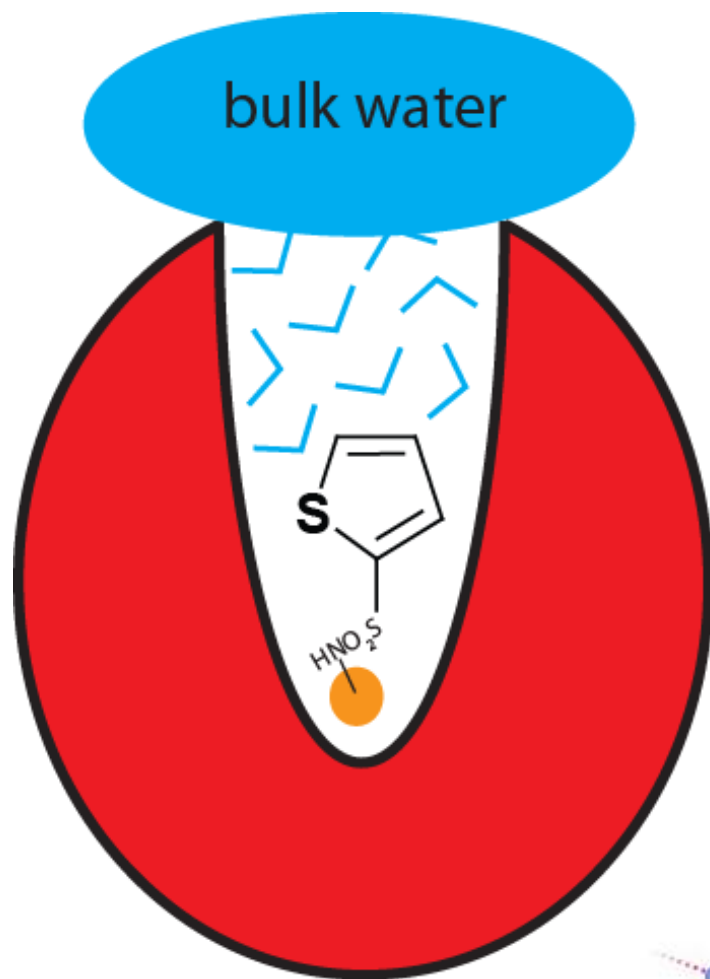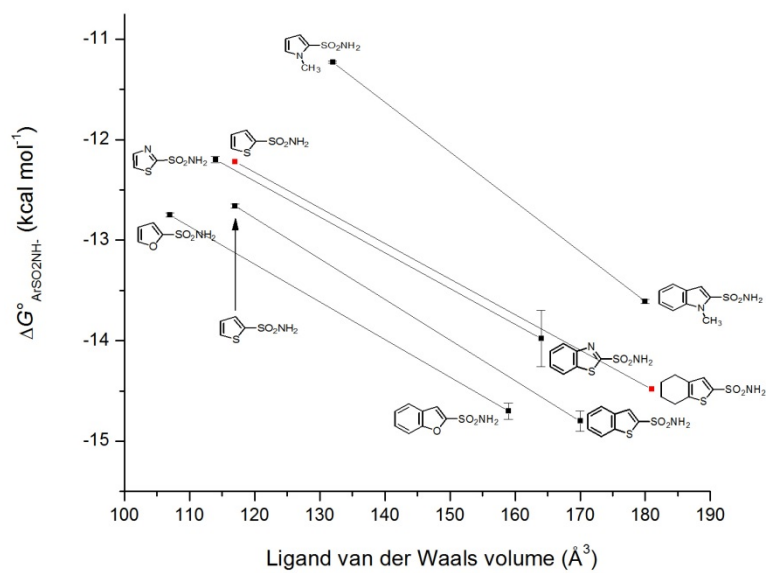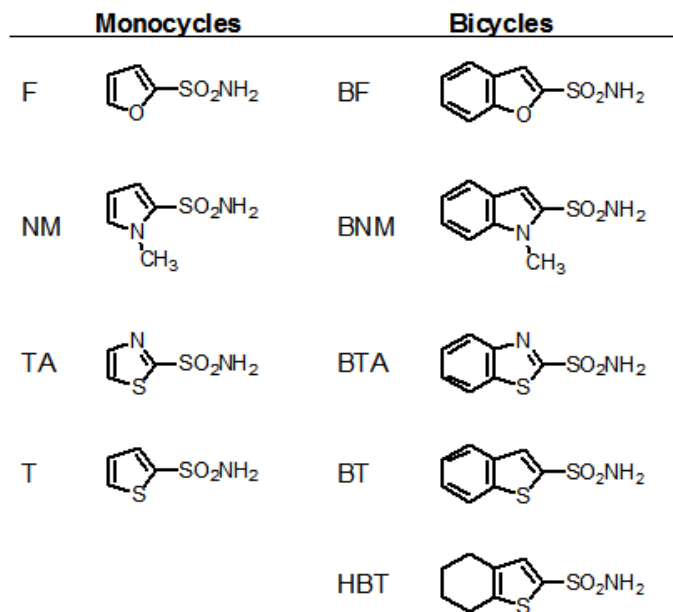
ΔΔG° ~ -25 cal Å⁻²
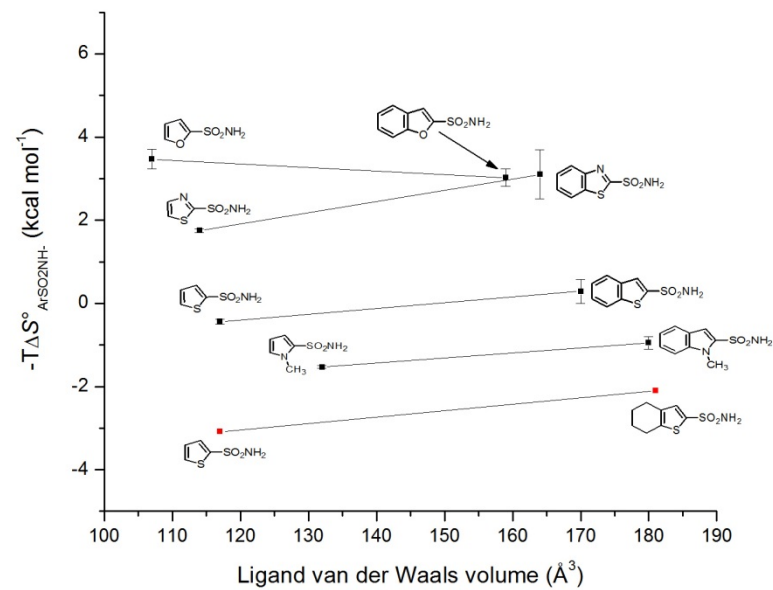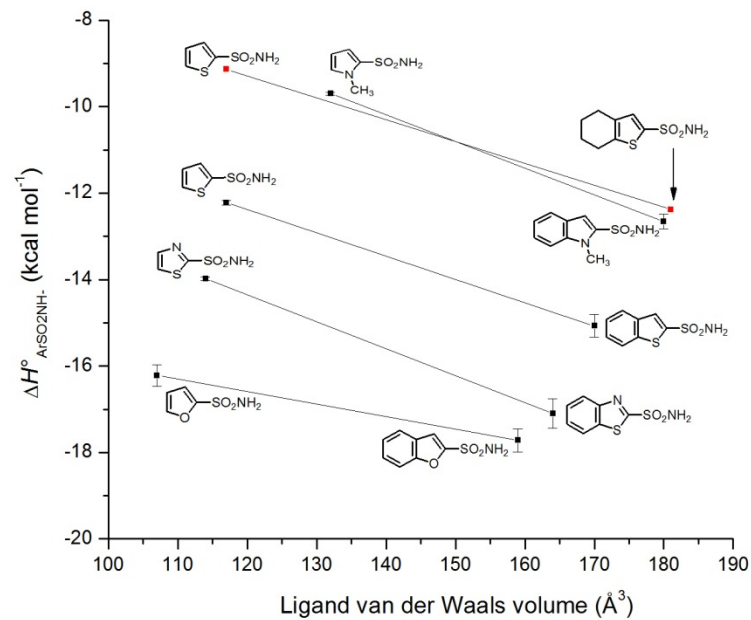
Opposite of hypothesized effect!
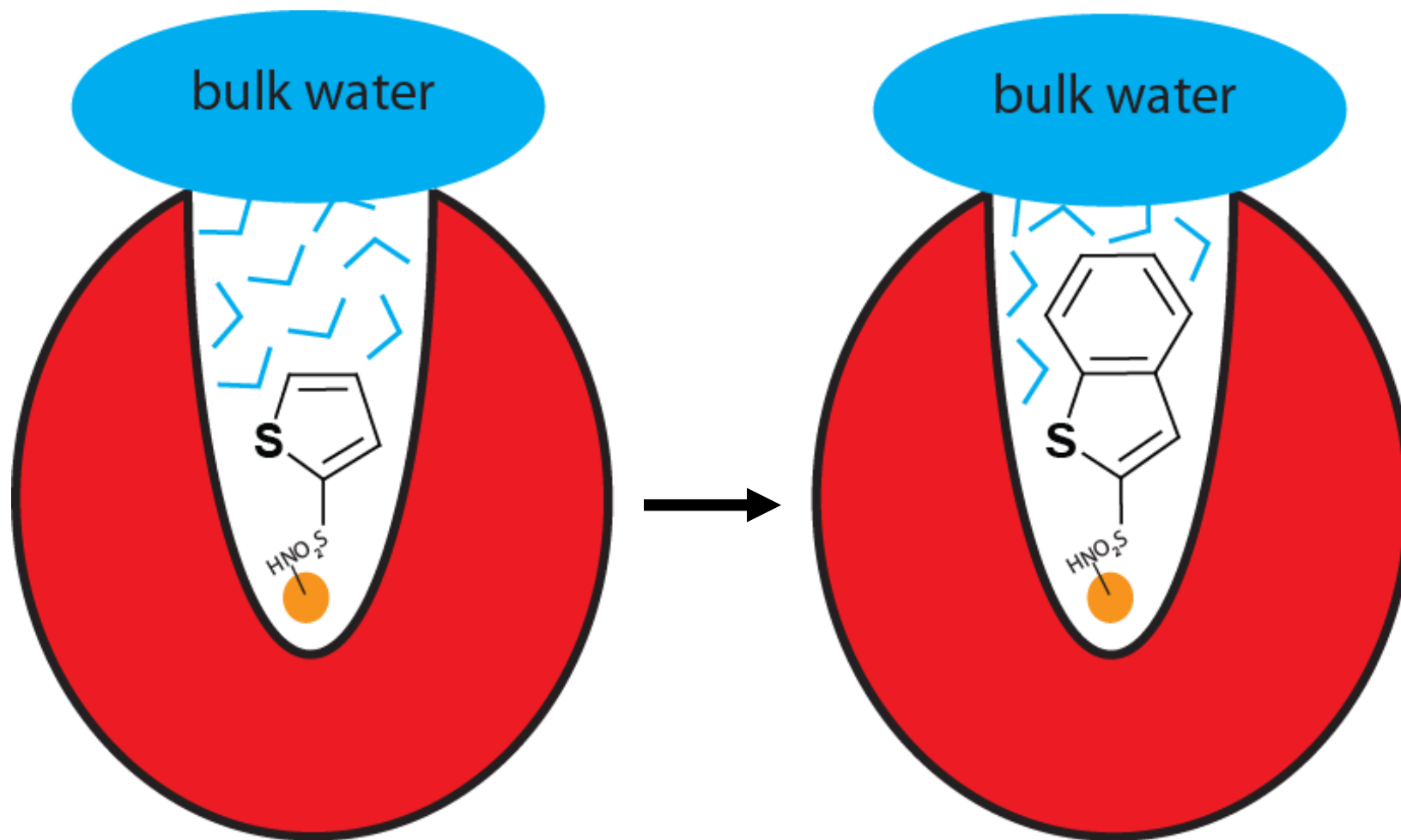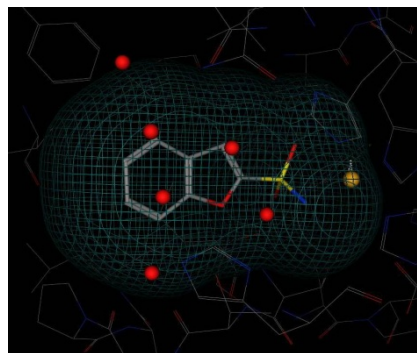
Enthalpy *not* Entropy drives the increase in affinity.
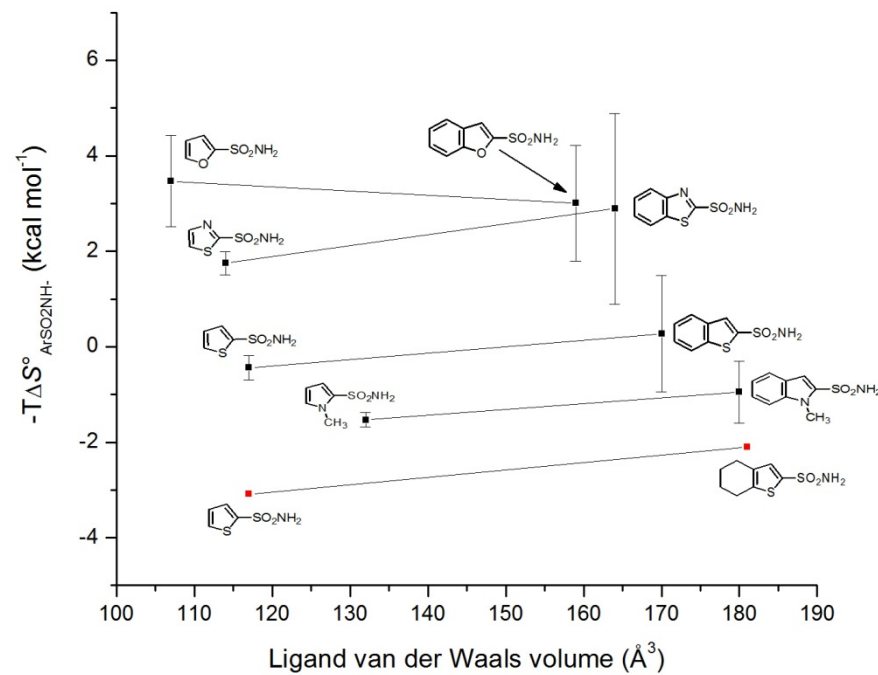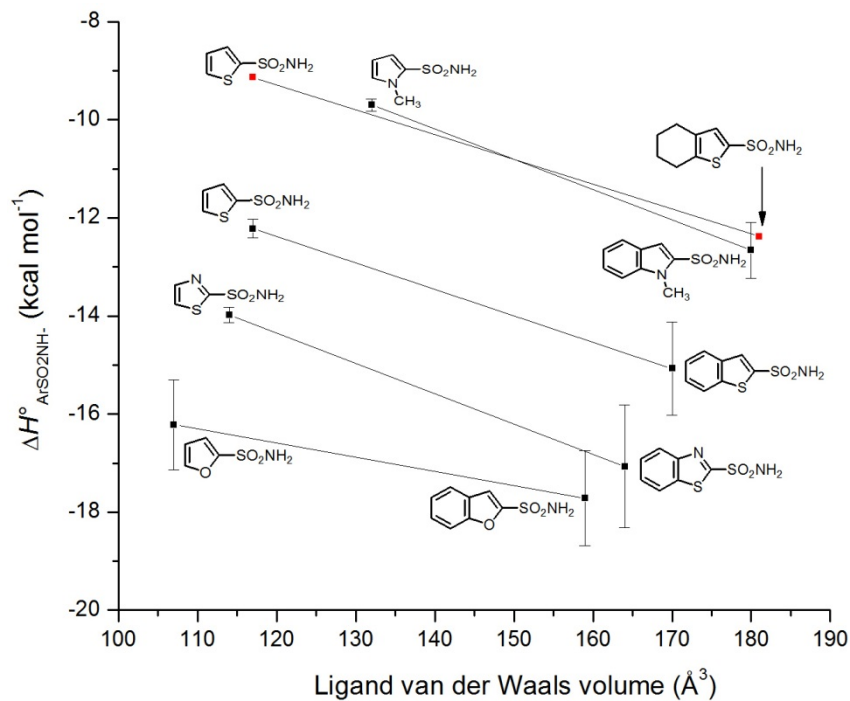
Same geometry
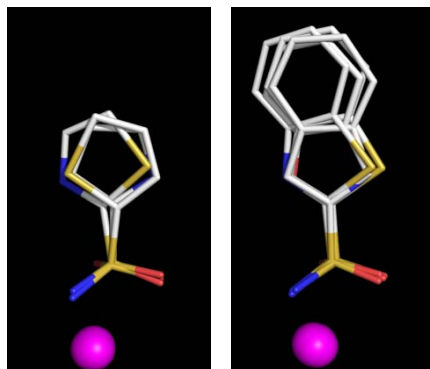
Active site waters

Interactions w/ HP wall
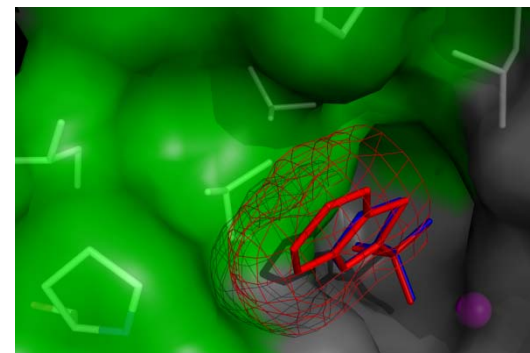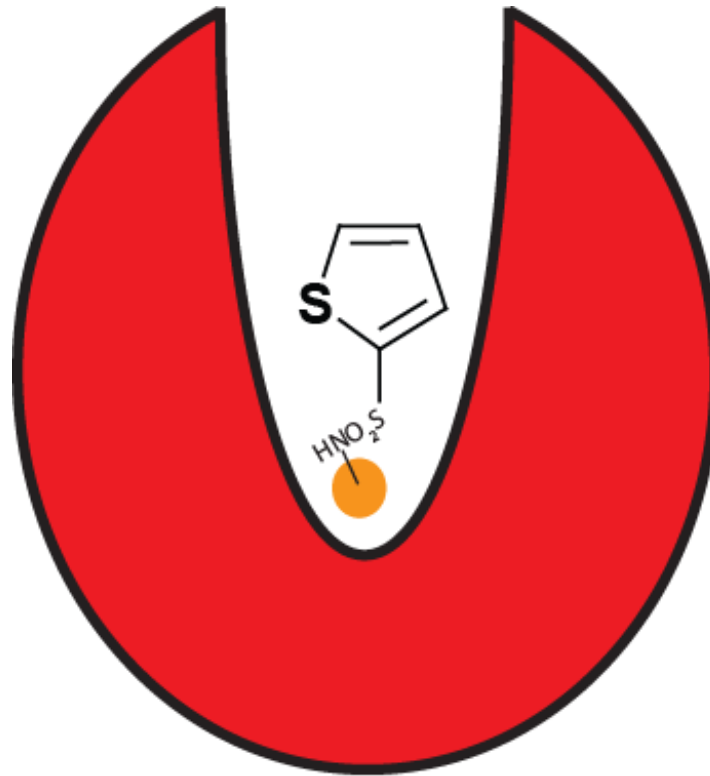
Modelin **ractions**

Add a "cap" of water molecules
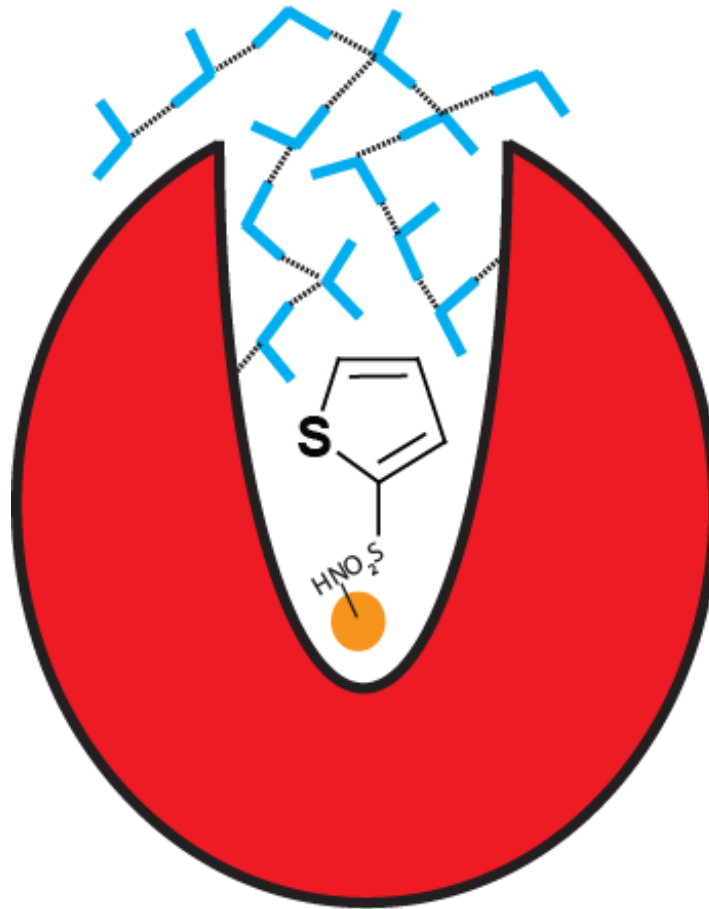
Ligand is highly solvent-exposed
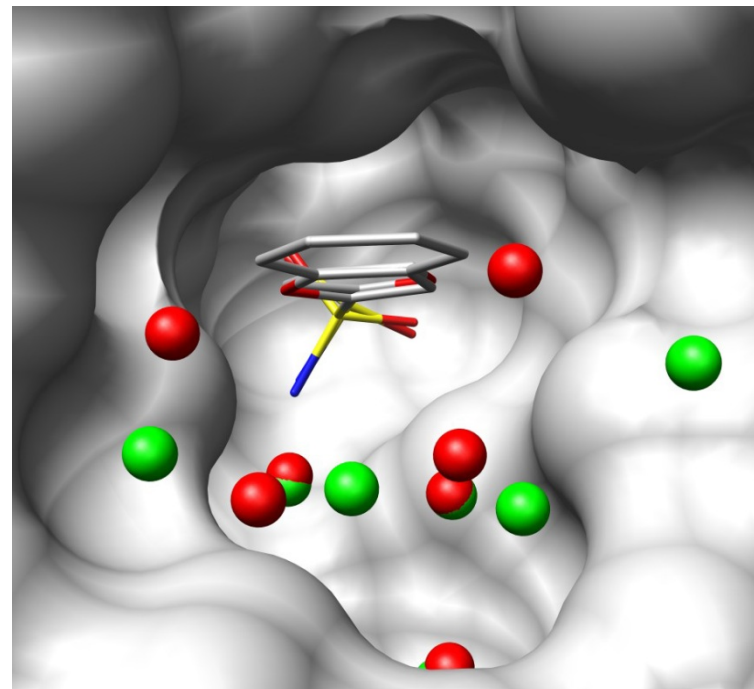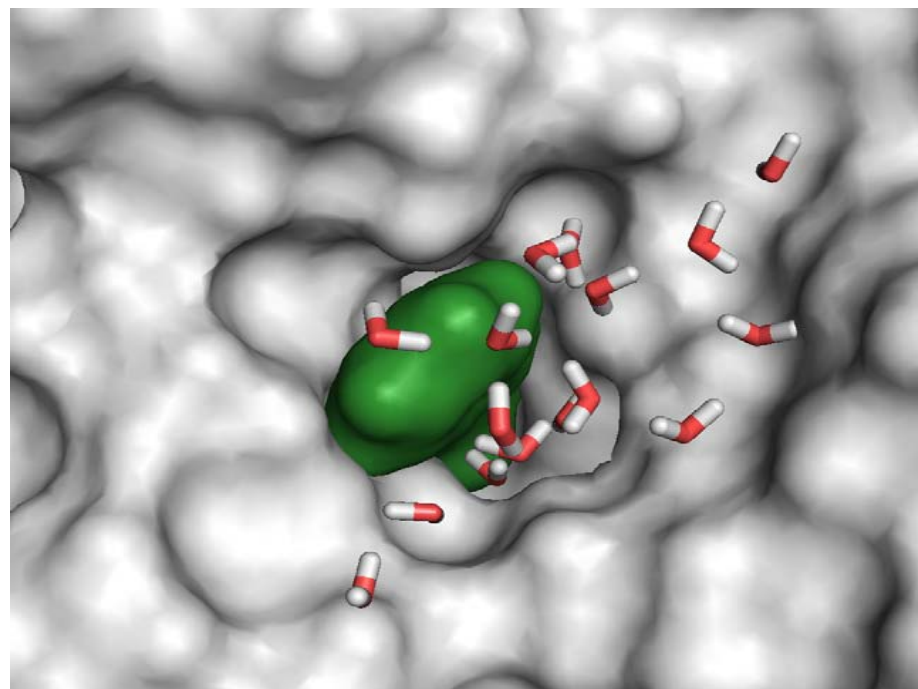
Fused ring appears to affect crystallographic water positions

Could the observed thermodynamic effect be solvent-mediated?
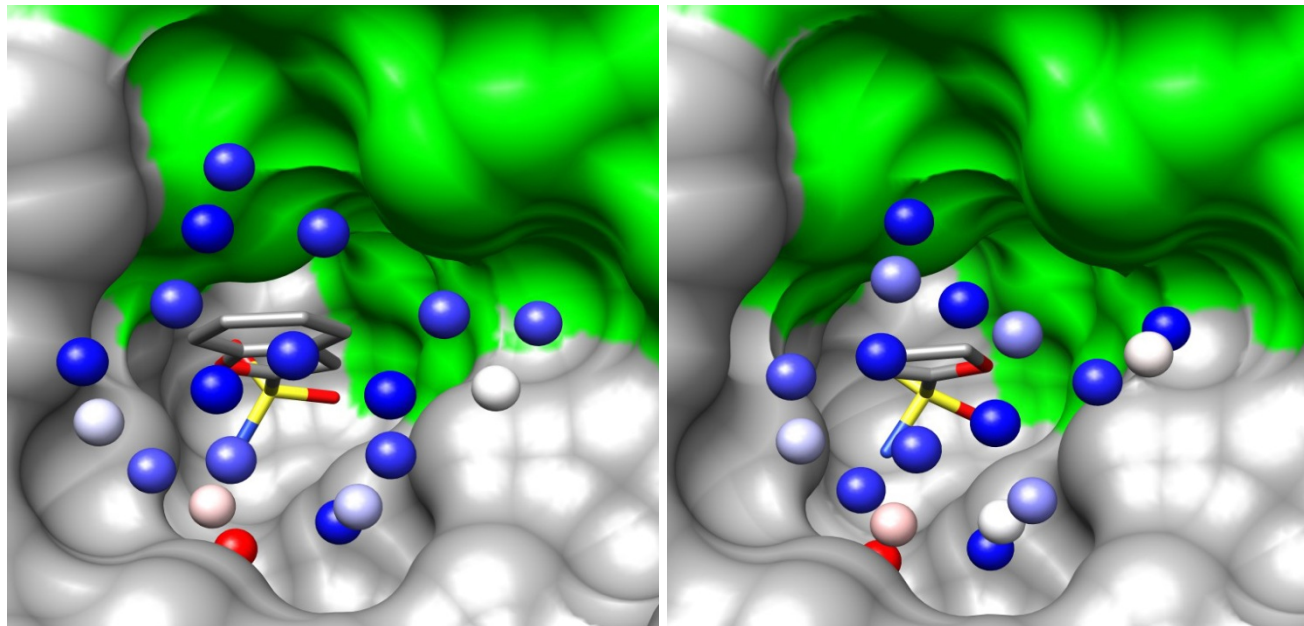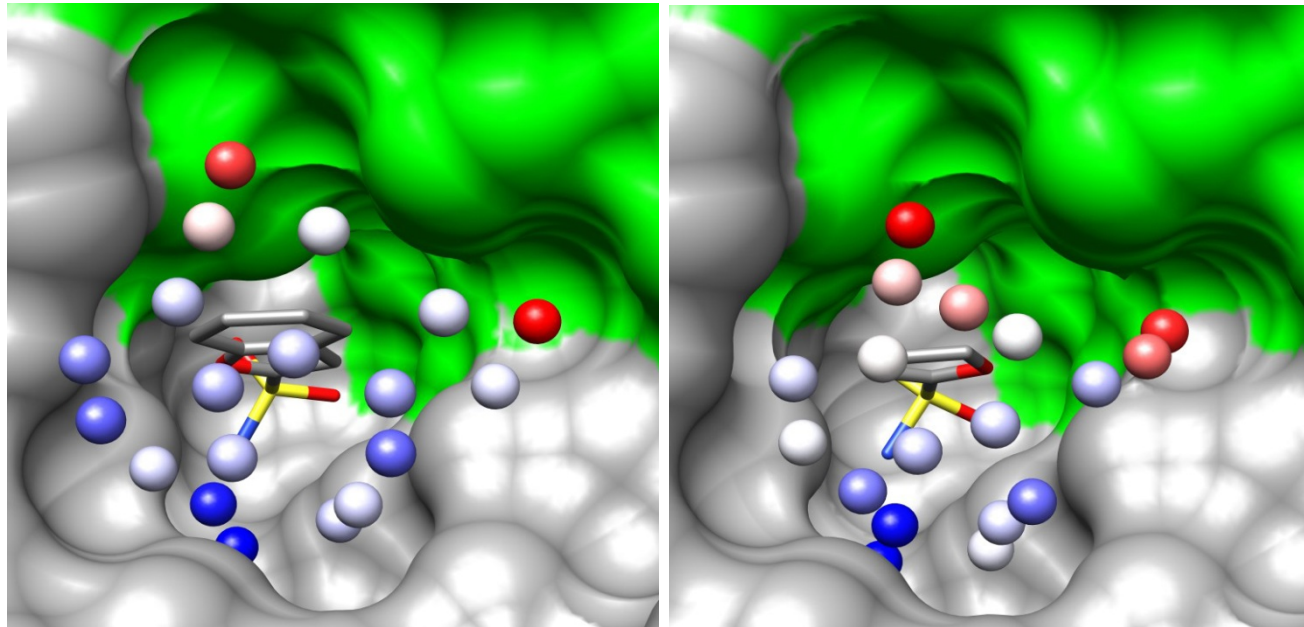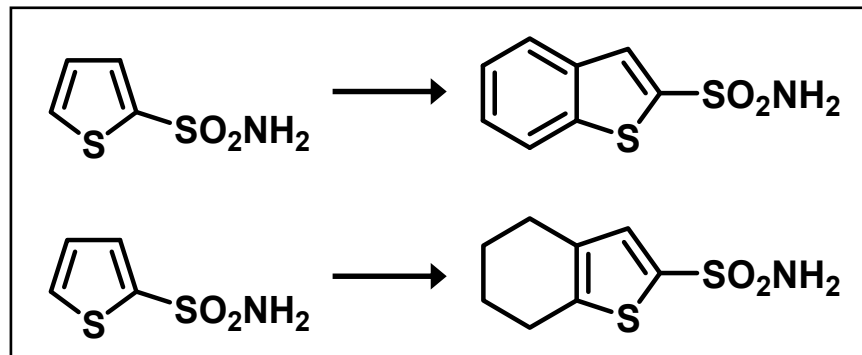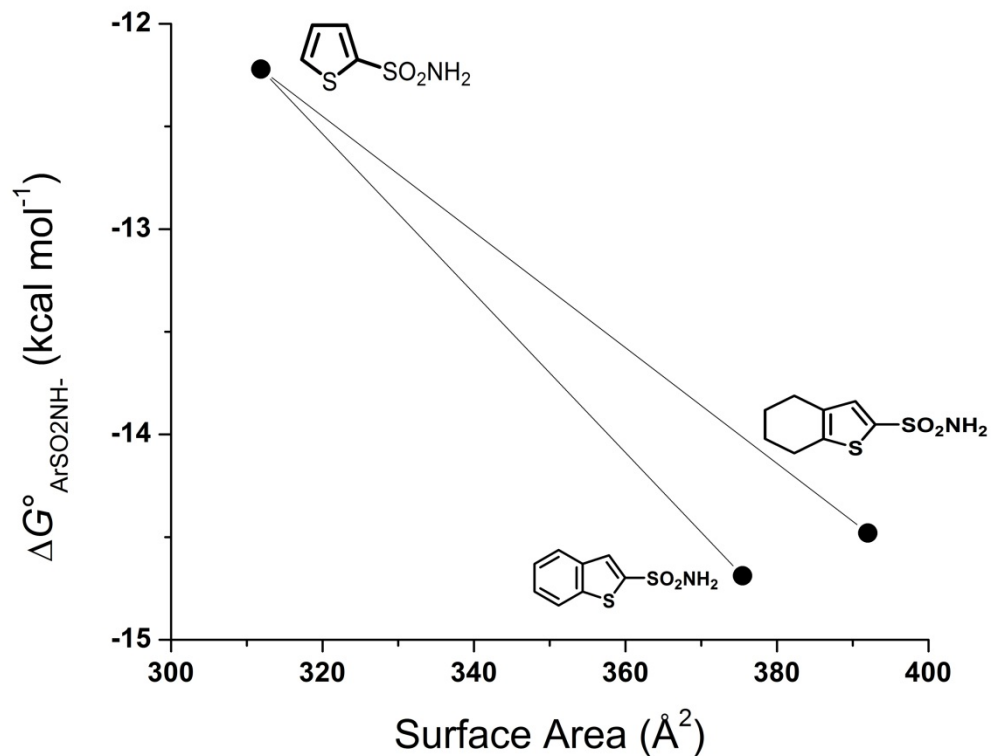
Results

1. The fused benzo ring makes waters in contact with hydrophobic residues less enthalpically unfavorable

2. The waters trapped between the fused benzo ring and the polar residues become more enthalpically favorable

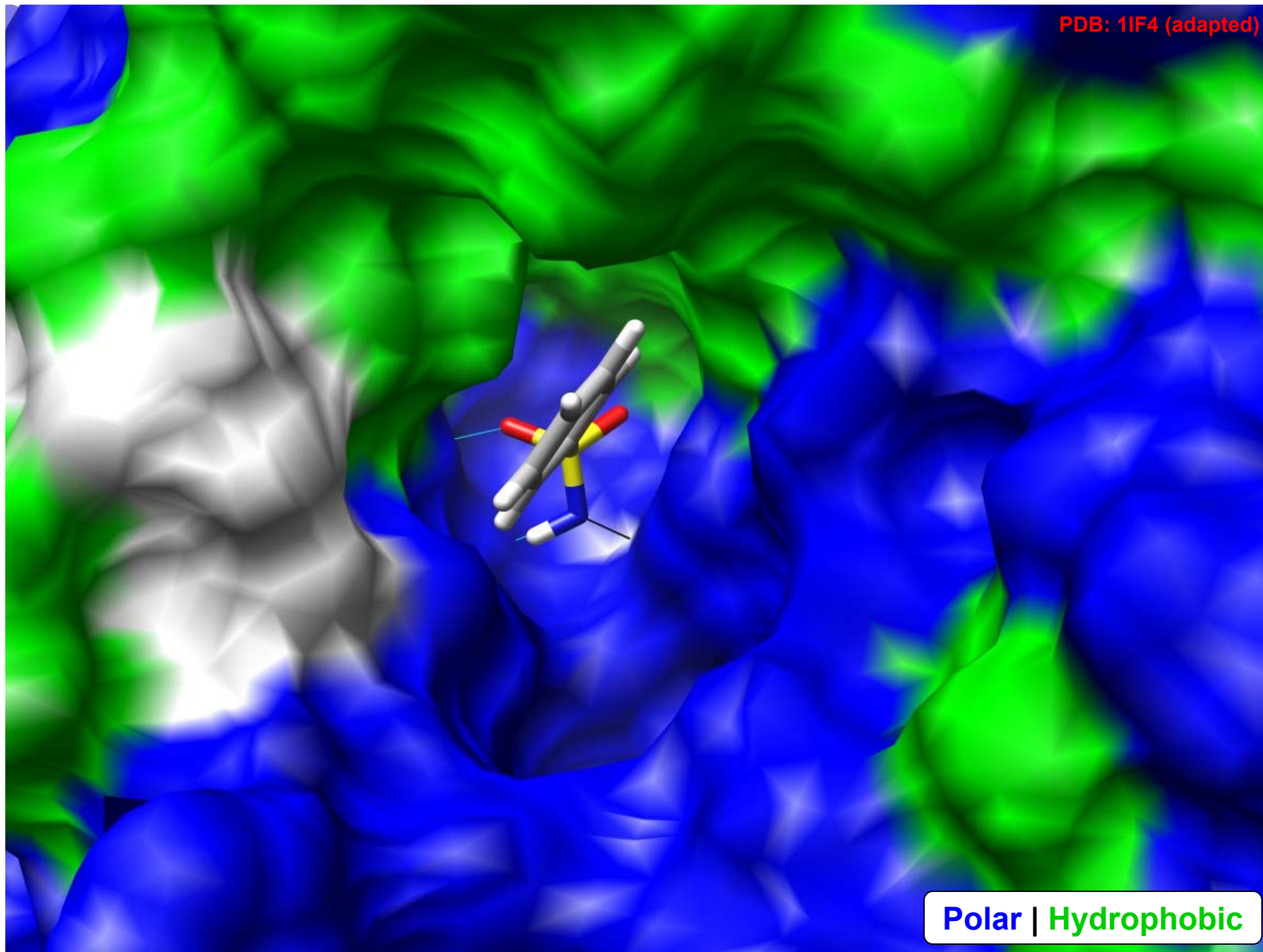3. The entropies of the waters solvating both F and BF ligands are similar

Aryl & alkyl: similar effects on ΔG°

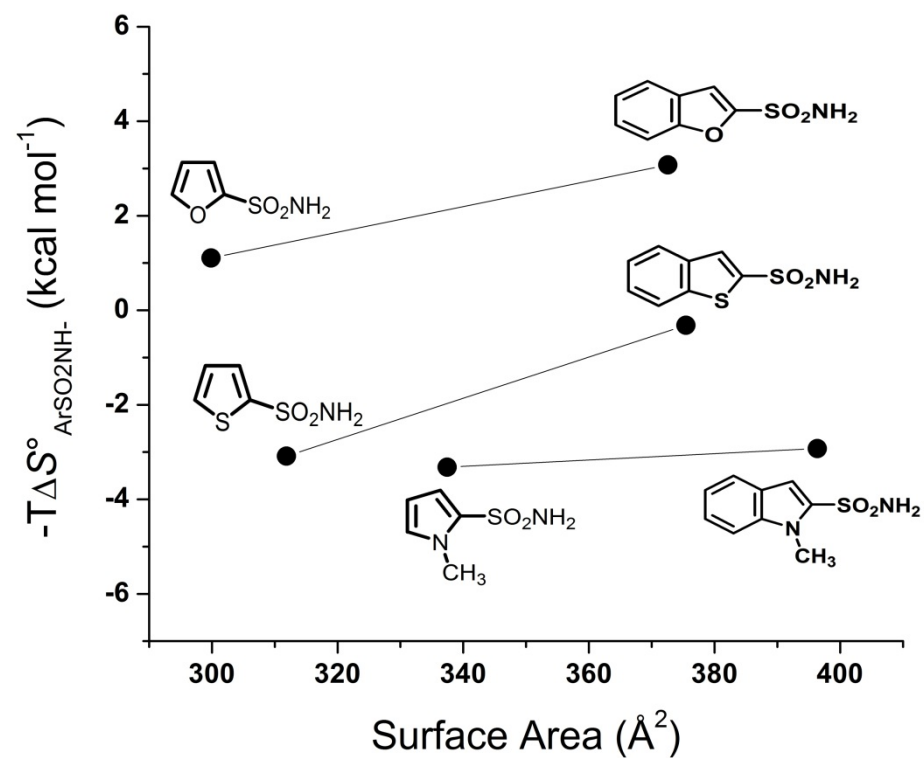# So: What *is* the Hydrophobic Effect?

- It is *not* the apposition of two non-polar surfaces with release of hydrogen-bond networks.

- It might be:
  - Some water release
  - Some restructuring of hydrogen-bond networks

Polar | Hydrophobic

ΔΔH° and TΔΔS° are unexpected
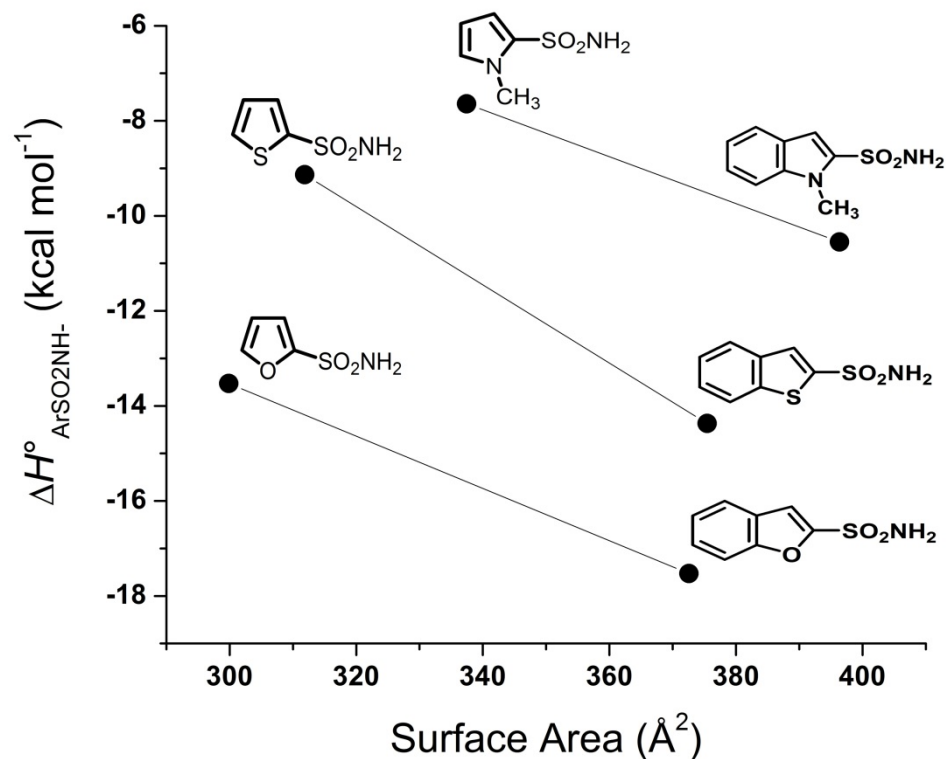
Opposite of hypothesized effect

Hit / Lead to Drug
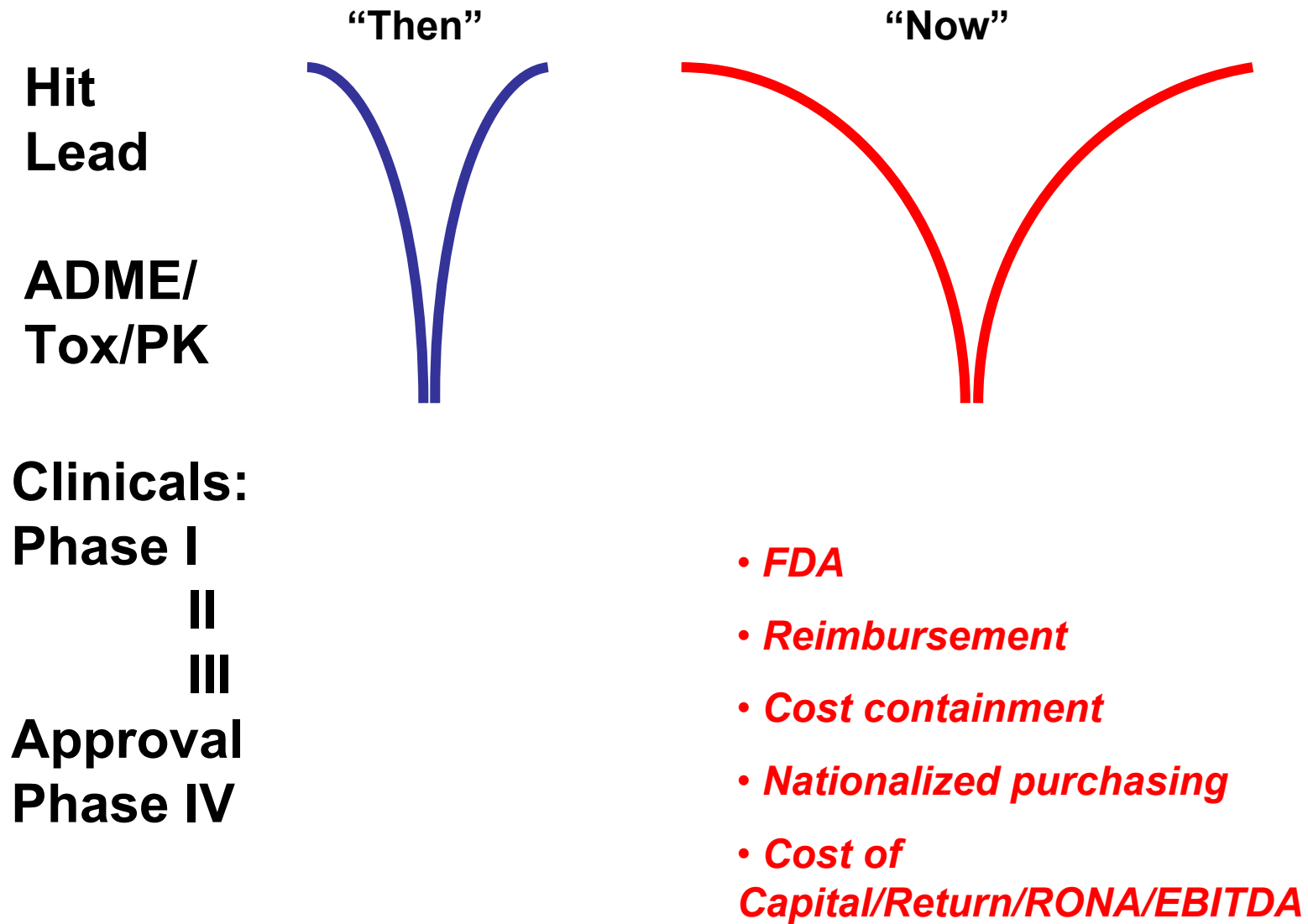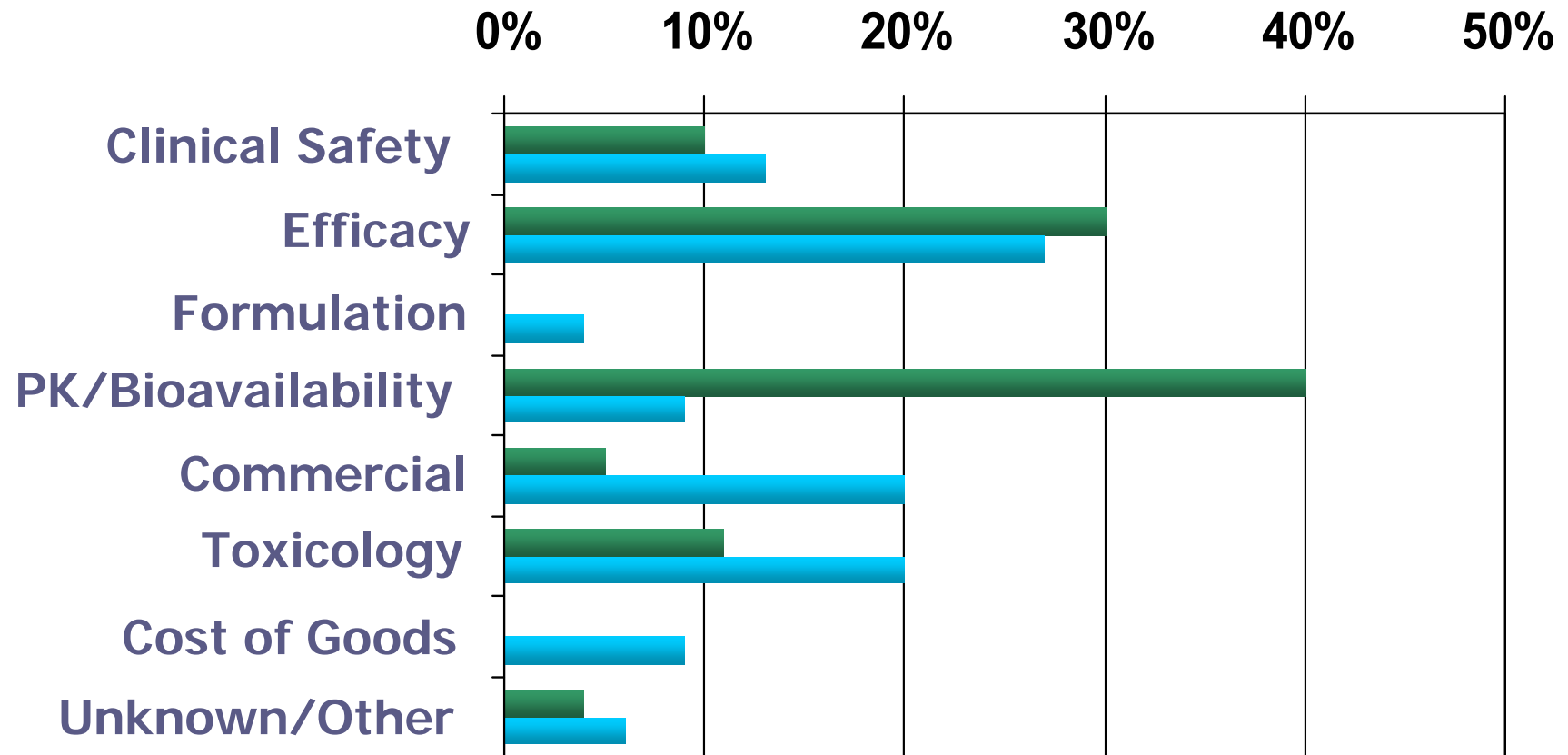
"Then" **"Now"**

**Hit
Lead**

**ADME/
Tox/PK**

**Clinicals:
Phase I
        II
        III
Approval
Phase IV**

- *FDA*

- *Reimbursement*

- *Cost containment*

- *Nationalized purchasing*

- *Cost of Capital/Return/RONA/EBITDA*

Reasons for attrition

Enthalpic and Entropic Contributions to the Free Energy of Interaction of Two Ions
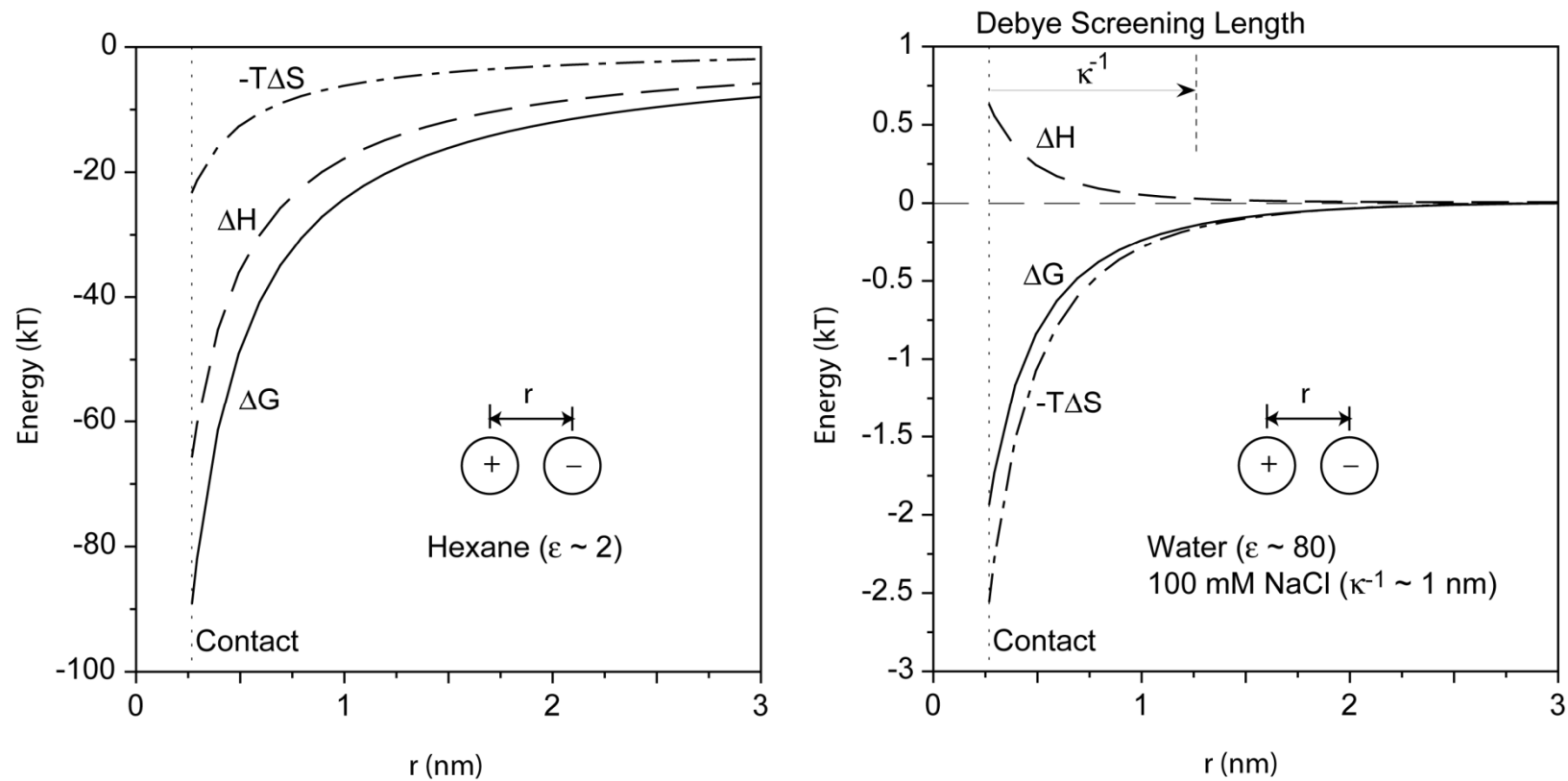
In hexane: enthalpically driven
In aqueous solution with 100 mM salt: entropically driven
    (related to the temperature dependence of ε_{H_2O})

Enthalpic and Entropic Contributions to the Free Energy of Interaction of Two Ions in Water

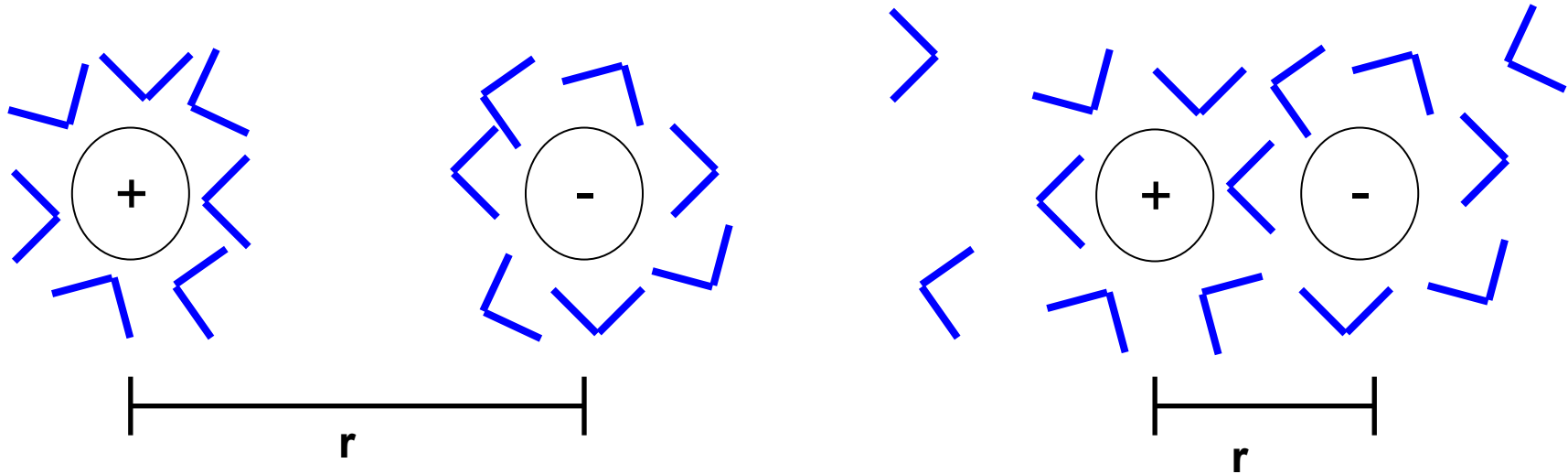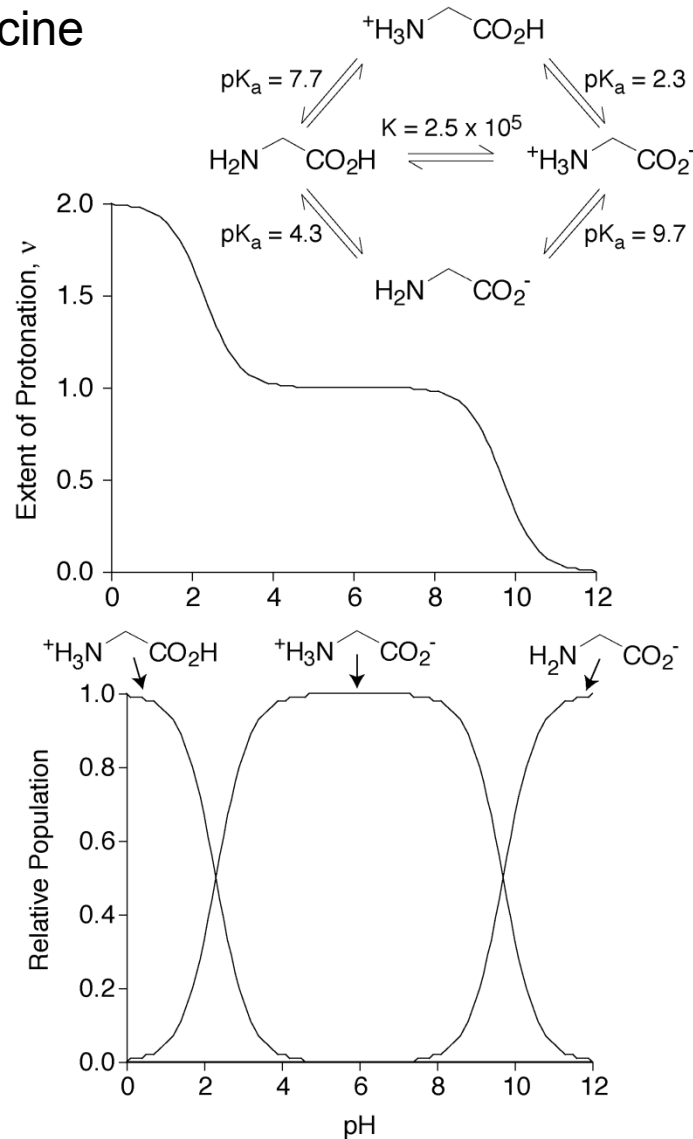- $\Sigma \Delta H_{ion-water} > \Delta H_{ion-ion} \Rightarrow \Delta H_{association} > 0$ (unfavorable) Temperature dependence of ε is a measure of the strength of ion-dipole interaction
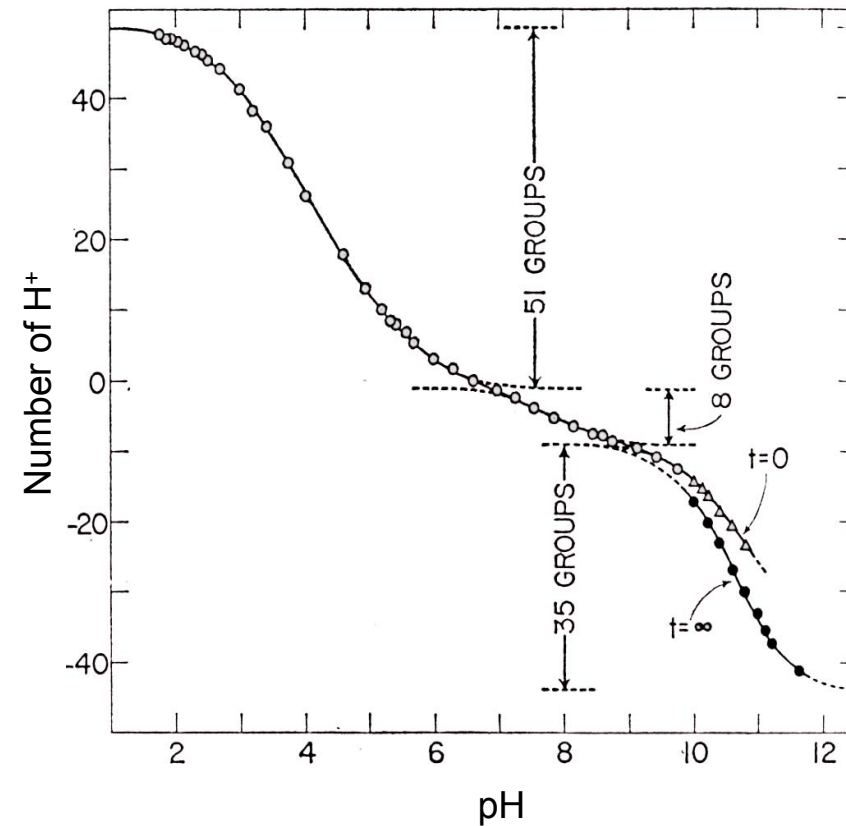- $\Delta S_{association} > 0$ (favorable) due to solvent release

Cooperativity in Proton Binding: Charge Regulation

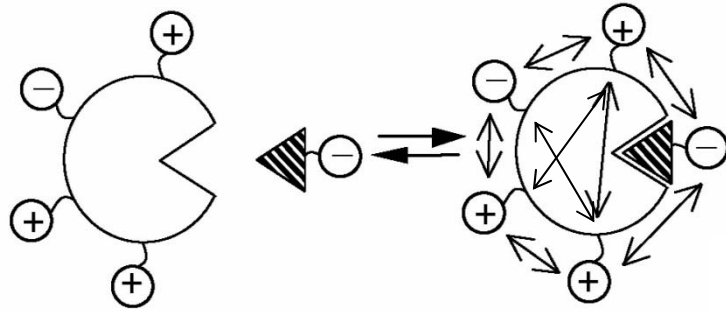Ionization constant of a group is influenced by the charge states of neighboring groups

Glycine

β-lactoglobulin at I = 0.15 M[1]

[1]Tanford, C. *Adv. Prot. Chem.* **1962**, 62, 69-165

Electrostatics in Proteins: Network of Charges

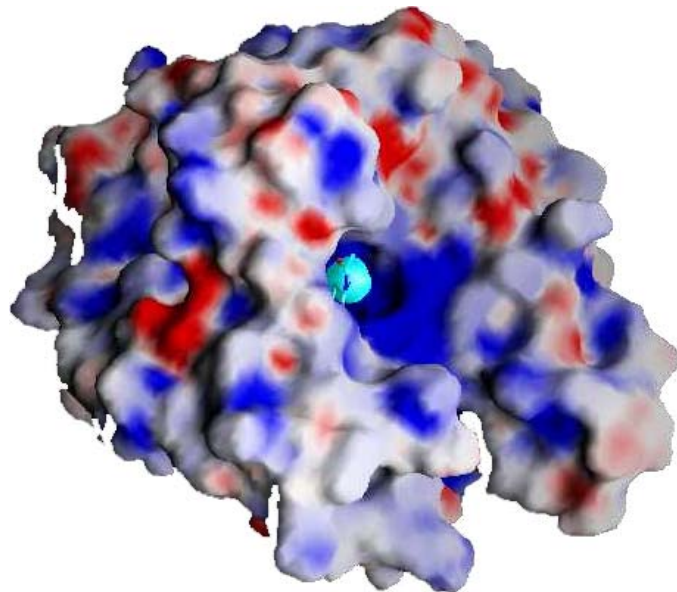
Charge Compensation: ionization state of one group alters the ionization state of another via "local pH" or "pK_a"

Regions of multiple dielectrics:
   ($\varepsilon(H_2O) = 80$; ε(protein core) = 2-5;
      ε(boundary layer H_2O) = 10-15)

Conformational changes/binding of ligand alters the dielectric cavity
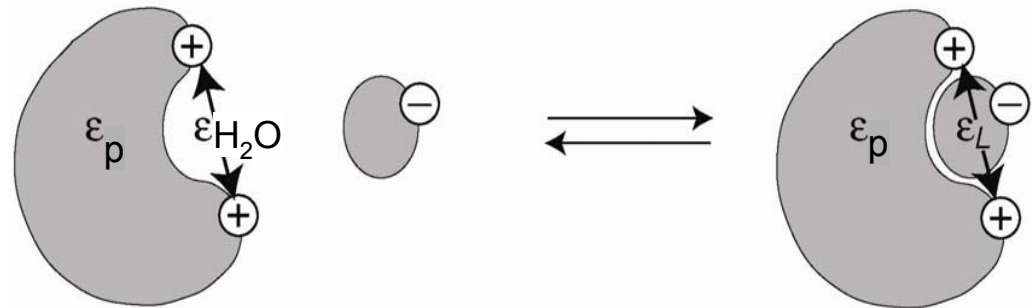
Electrostatic map of HCA I with sulfonamide ligand

Protein Plasticity: Ligand-Induced Conformational Change of Receptor
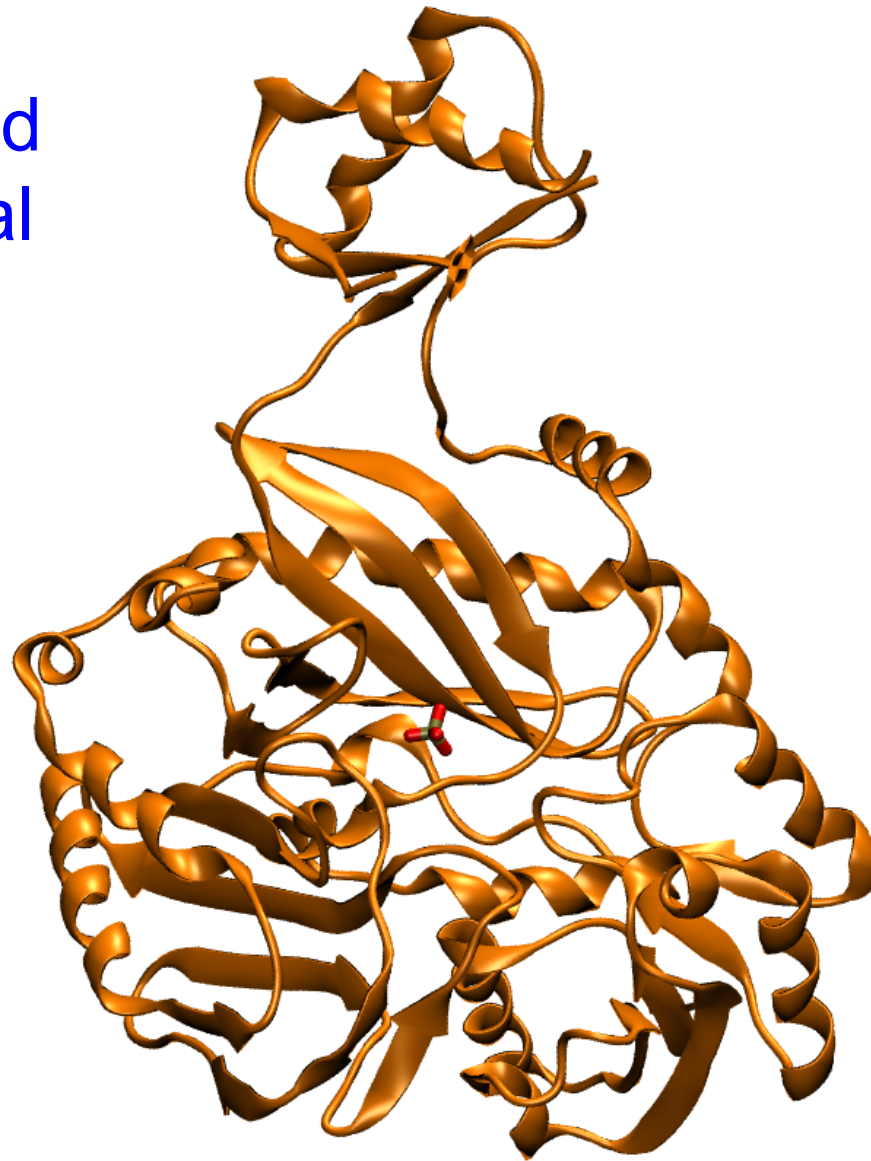
Unliganded *E. coli* biotin carboxylase. PDB code 1BNC

Ligand Induced Conformational Change of Receptor

E. coli biotin carboxylase bound to ATP. PDB code 1DV2

Ligand Induced Conformational Change of Receptor

E. coli biotin carboxylase unliganded (left) and bound to ATP (right)
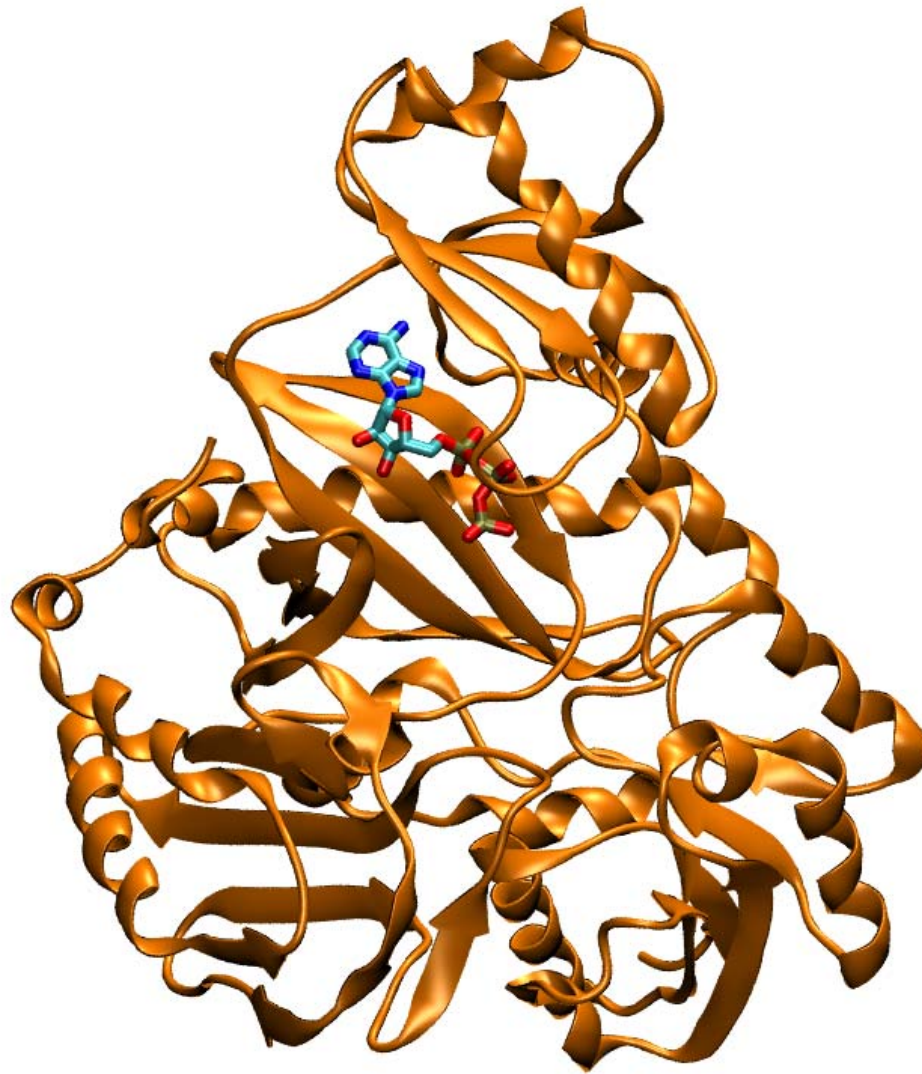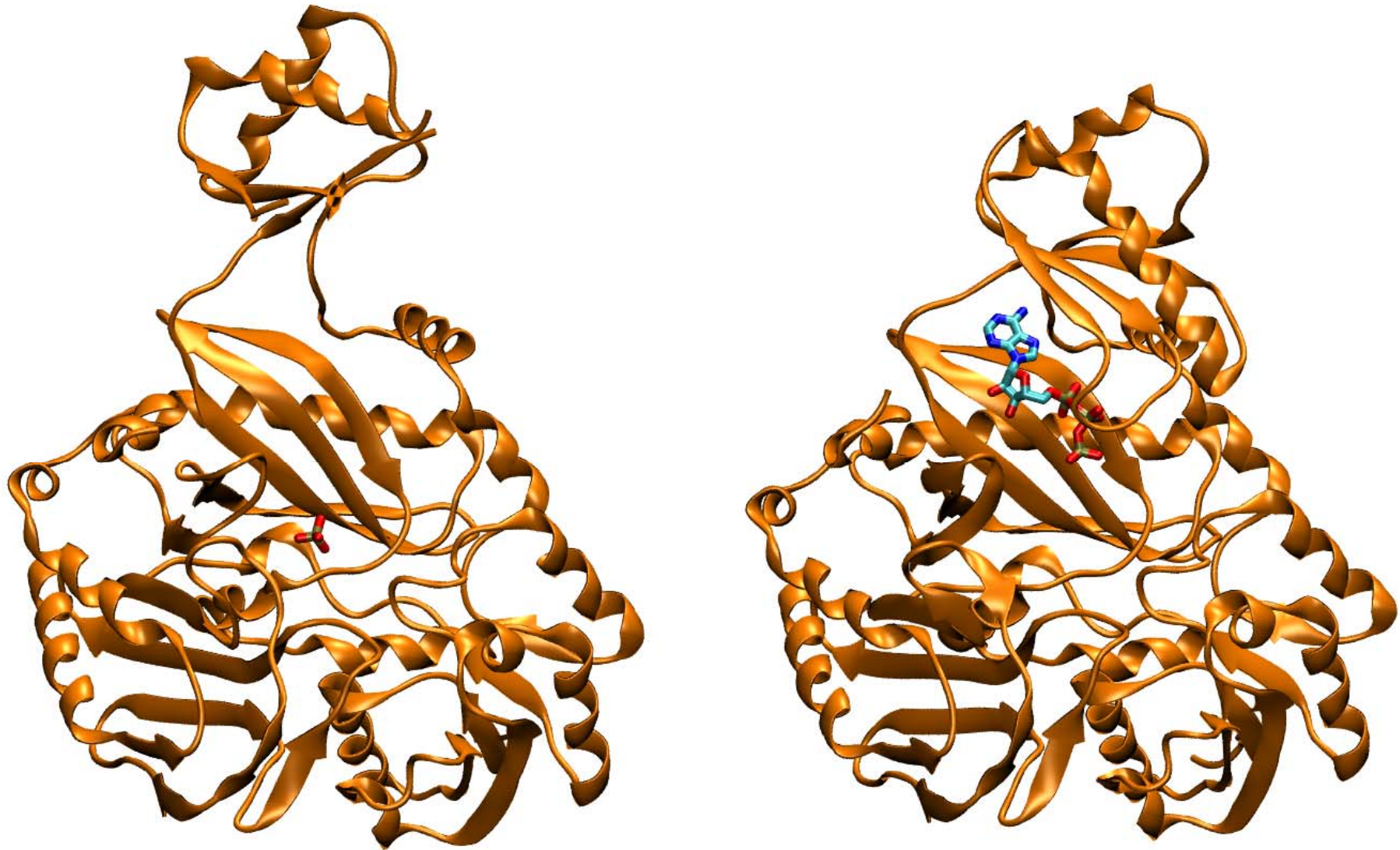PDB codes 1BNC, 1DV2

Plasticity in Proteins

Cooperativity is observed between residues in DHFR separated by large distances.

In *E. coli* DHFR, mutations of Gly-121 and Met-42 have a synergistic effect upon enzyme catalysis.

Effect of mutations upon enzyme kinetics

	Gly-121	Gly-121-Val
Met-42	100%	0.5%
Met-42-Phe	67%	0.34% (expected) **0.017% (measured)**

20 Å

Benkovic, S.J. and Hammes-Schiffer, S. *Science* **2003**, 301, 1196-202.
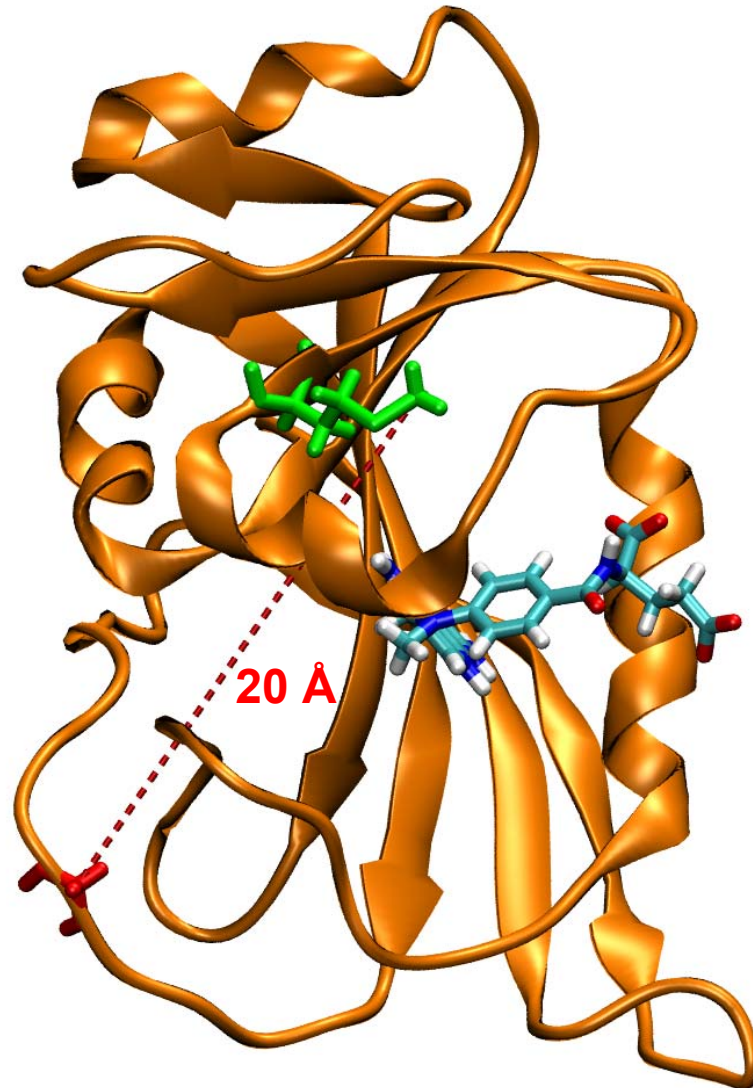
Enthalpy/Entropy Compensation: Theoretical Model

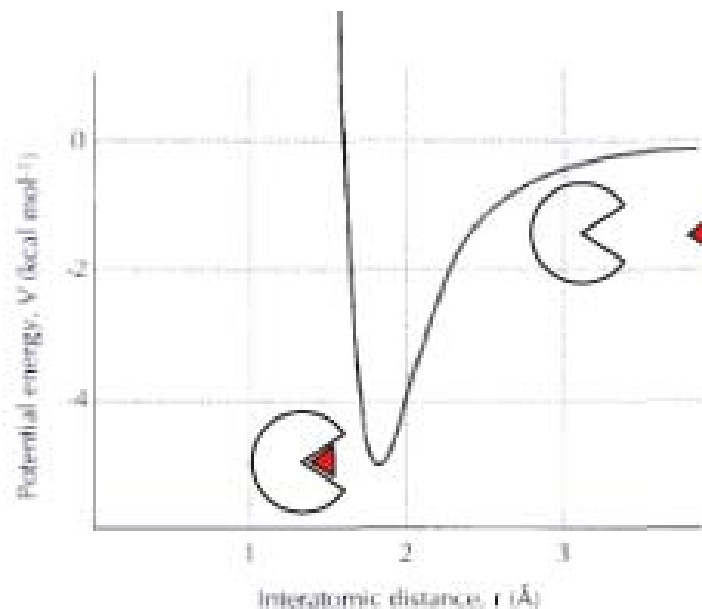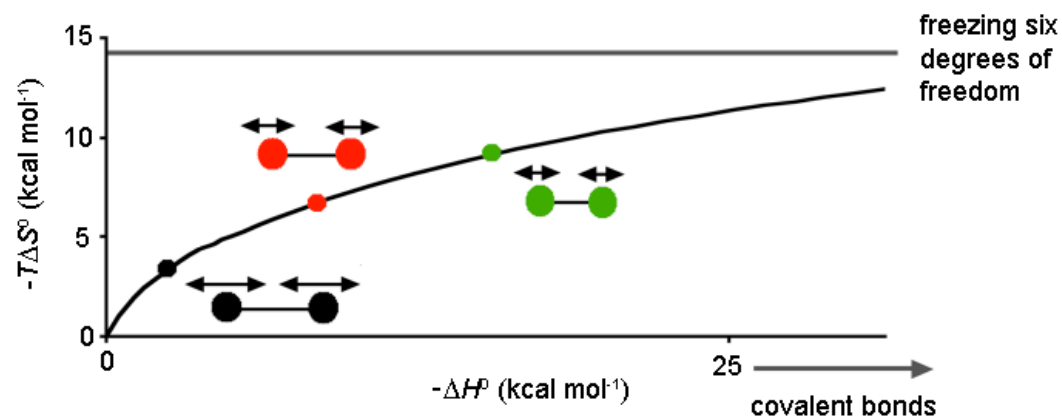- Model treats residual mobility of ligand-protein complex as "spring"

- Model relates $T\Delta S^\circ$ and ΔH°

Dunitz, J. *Chem. Biol.* **1995**, *2*, 709-712.
Williams, D.H. and co-workers *Angew. Chem. Int. Ed.* **2004**, *43*, 6596-6616.

The Zeta Potential of Polyethylene

- **Observation:** the surface of polyethylene in contact with neutral water is negatively charged.

- **Interpretation:** OH- associates preferentially with non-polar interfaces.

- **Relevance:** Much/most of molecular recognition is the hydrophobic effect.

- **Origin:** Who knows?
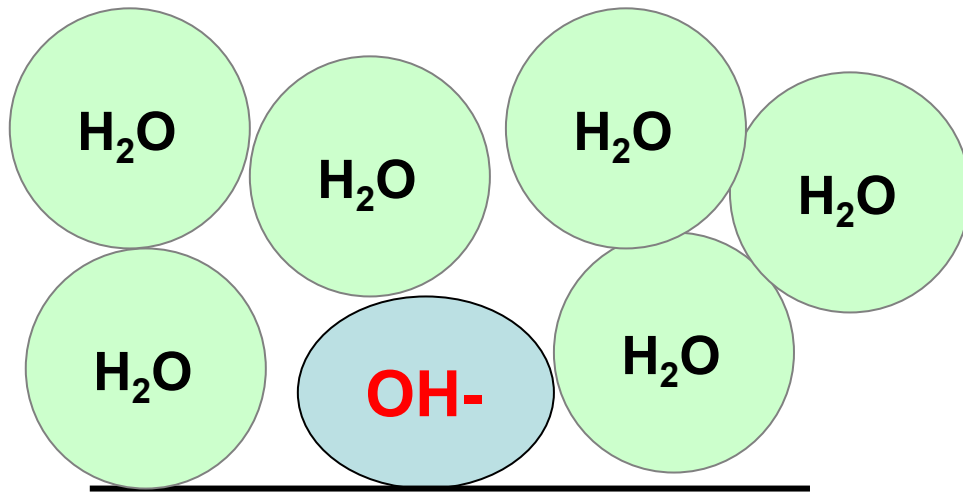
Polyethylene (or any other low

Dielectric constant matter (liquid, solid, vapor)

- Charge in a dielectric cavity?

- Hydrogen-bond Network?

- Enthalpy/Entropy?

Carbonic Anhydrase: A Model Protein

- Commonly used model protein for physical-organic studies

- Stable (T_m = 65 °C)

- Monomeric, 30 kDa

- No disulfide bonds

- Structure is dominated by 10 β-sheets

- Zn(II)-OH cofactor in active site

- Function: CO_2 hydration

- Binding of sulfonamide inhibitors is well-characterized